Farsighted Rationality in Hedonic Games

G.-Herman Demeze-Jouatsa*[1] and Dominik Karos†[2]

[1] *Bielefeld University, Center for Mathematical Economics*
[2] *Bielefeld University, Center for Mathematical Economics, Chair of Economic Theory*

October 27, 2021

Abstract We consider a hedonic coalition formation game in which at each possible partition any new coalition can decide the probability with which to form and leave the current partition. These probabilities are commonly known so that farsighted players can decide whether or not to support a coalition's move: they know which future partition, and hence payoffs, will be reached with what probability. We show that if coalitions make mistakes with positive probability, i.e., if they choose probabilities that are always above some $\varepsilon > 0$, then there is a behavior profile in which no coalition has a profitable one-shot deviation.

Keywords: abstract games, hedonic games, farsighted stability, coalition stable equilibrium

JEL: C71, C72

1 Introduction

An abstract game consists of a set of states (or outcomes), agents' payoffs in each state, and an effectivity correspondence that describes for any two state what coalitions are able to implement a move from the former to the latter. Because of their generality, abstract games can be used to model a great variety of games; in particular, games with non-transferable utilities: there, a state comprises a partition of players into coalitions and a payoff for each player. The class of games that will be interesting

*G.-Herman Demeze-Jouatsa gratefully acknowledges financial support from the DFG (Deutsche Forschungsgemeinschaft / German Research Foundation) via grant Ri 1128-9-1 (Open Research Area in the Social Sciences, Ambiguity in Dynamic Environments)

†corresponding author. dominik.karos@uni-bielefeld.de

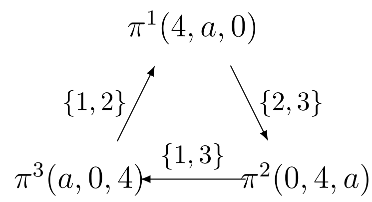

Figure 1: The roommate problem

in this paper are *hedonic games*, which were introduced by Drèze and Greenberg (1980). Here the idea is that whenever some coalition S forms, each player's payoff is predetermined. That is, while coalitions might be in competition with each other, there is no intra-coalition competition about resources or payoffs. Among others these games include matching problems or network formation games. One very well known example of a hedonic game is the 'roommate problem' that is depicted in Figure 1. There are three players who have to decide about who will be moving in together, i.e., about how to form a partition. Denote by π^1, π^2, and π^3 the partition in which players 1 and 2, 2 and 3, or 1 and 3, respectively are roommates, while the remaining player is excluded. From Figure 1 we observe that for $a \in (0, 4)$ preferences are as follows: 1 prefers to move in with 2 over moving in with 3 over staying alone; 2 prefers moving in with 3 over moving in with 1 over staying alone; and 3 prefers moving in with 1 over moving in with 2 over staying alone. Once a partition has formed, there is no more negotiating about payoffs, but everything is fixed. Unfortunately, despite their rather simple structure, hedonic games are not easily solved. Bogomolnaia and Jackson (2002) and Banerjee et al. (2001) provide sufficient conditions for the nonemptiness of the core; Iehlé (2007) provides a condition which is both necessary and sufficient and very similar to the balancedness condition by Shapley (1967) and Bondareva (1963). The game in Figure 1, however, does not obtain a core stable partition.

More recently, other solutions to cooperative games in general, and to hedonic games in particular, have gained some attention: namely, farsighted solutions that first emerged from Harsanyi (1974) and Chwe (1994) and were applied to hedonic games for instance by Diamantoudi and Xue (2003). The general idea behind farsighted solutions is that coalitions do presume to remain in the state they deviate to but acknowledge and expect other coalitions to react and immediately leave the new state. When using farsighted solutions in the context of coalition formation games there are some structural obstacles to overcome: first, unlike in the myopic case, it is essential if and how players react who have been left behind by a moving coalition as this will affect future moves. Second, when deciding whether or not to leave a state coalitions should not only consider the long-term effect of their moving, but also the long-term effect of their not moving. The first point was taken up by Ray and Vohra (2015) who proposed conditions on the structure of an effectivity correspondence that emerges from a coalition formation game; the second point was the topic of Karos and Robles (2021) who provided a solution based on expectation functions (cf. Dutta and Vohra, 2017) that allowed coalitions to form expectations about the future in case they remain in the status quo. While these expectation functions have clear axiomatic and non-cooperative foundations, their big disadvantage is that they may not exist for some games. For instance, they do not exist for the roommate problem depicted in Figure 1.

In this paper we shall consider hedonic games and extend the idea of Karos and Robles (2021) by using their non-cooperative foundation and allow coalitions to play mixed strategies. The idea is simple enough: provided that payoffs are sufficiently well-behaved we might hope for the existence an equilibrium in mixed strategies. And we will indeed find some positive result. Mathematically speaking we proceed as follows: we define for each coalition at each partition a probability distribution over states it might move to, a so called *mixed coalition behavior*. These strategy profiles define transition probabilities among states, and if we restrict ourselves to strictly positive distributions, i.e., completely mixed behaviors, then these transition probabilities define an irreducible Markov process. As our state space consists only of partitions of the player set (recall that in hedonic games players' payoffs are uniquely determined for every partition), the Markov process is recurrent as well, which means that it obtains a unique stationary distribution. This distribution describes how much time the process will (in average) spend in each state, i.e., how long each partition will hold, so that we can define a player's utility as the expected utility over all partitions weighted by the stationary distribution. Thus, we define a coalitional game that specifies for each (completely mixed) strategy profile a payoff vector. We then turn to deviations and make the following observation, which is our technical main result: for any two irreducible finite space Markov processes whose transition matrices are identical everywhere but in one row, the stationary distribution of any convex combination of the two is a convex combination of the two stationary distributions of those processes. Observing that in hedonic games a coalition has only two options at any state, namely to form or not to form, and that a change of strategy at only one partition leads to a new Markov process that differs only in one row, reveals that for any fixed strategy profile of their opponents the feasible utility vectors of a coalition form a (bounded) line. In particular, when moving along this line players' payoffs will either increase or decrease, so that either all players agree that one end point of the line is better than the other, or no two points can be Pareto-ranked. This allows us to conclude that the set of best-responses that a coalition has (in the sense that they are not Pareto-dominated) is a convex set. From here, the rest is pretty straightforward: we show that the best response correspondence satisfies the conditions of Kakutani's fixed point theorem and, thus, has a fixed point. Hence, there is a mixed strategy profile from which no coalition has a profitable one-shot deviation. For the roommate problem above Karos and Robles (2021) show that such a profile exists.

The remainder of the paper is structured as follows: In Section 2 we introduce the necessary notation, define probabilistic expectation functions, and derive expected utilities using some well-known results from the literature on Markov processes. In Section 3 we introduce hedonic games and provide a formulation in terms of effectivity correspondences. Section 4 introduces the non-cooperative setup in which we endow coalitions with strategies, and in Section 5 we show that for any slightly perturbed game (meaning that all strategies are played with small but positive probability) there is an equilibrium. The paper concludes in Section 6 with a brief discussion.

2 Preliminaries

2.1 Abstract Games

Let N be a finite set of players with $|N| \geq 3$. Subsets $S \subseteq N$ are called *coalitions*. For $S \subseteq N$ write 2^S for the set of subsets of S, and $P(S)$ for the set of nonempty subsets. A *partition* is a collection $\pi = \{S^1, \ldots, S^m\}$ of nonempty coalitions such that $\bigcup_{k=1}^m S^k = N$ and $S^k \cap S^l = \emptyset$ for all $k \neq l$. For $i \in N$ and a partition π we write $\pi(i)$ for the unique element of π that contains i. The set of all partitions is denoted by Π.

Let X be a finite set of states. An *abstract game* is a tupel $\big(N, X, E, (U_i(\cdot))_{i\in N}\big)$, where $U_i : X \to \mathbb{R}$ is player i's utility function over states and $E : X \times X \rightrightarrows 2^N$ is an *effectivity correspondence*: for two states $x, y \in X$ the (possibly empty) set $E(x, y)$ comprises all coalitions that are *effective* for a move from x to y, i.e., that can replace x with y. We assume that $E(x, x) = 2^N$, that is, each coalition can decide not to change the status quo; and $\emptyset \in E(x, y)$ if and only if $x = y$.

A lottery over X is a probability measure over X, the set of all lotteries over X is denoted by $\Delta(X)$. Players in the abstract game are expected utility maximizers, that is for any lottery $\lambda \in \Delta(X)$ their utility is given by $u_i(\lambda) = \sum_{x \in X} \lambda(x) U_i(x)$.

2.2 Probabilistic Expectation Functions

Let $\left(N, X, E, (U_i(\cdot))_{i \in N}\right)$ be an abstract game. One way to deal with farsighted deviations is by endowing coalitions with expectations about who is deviating where. This is the path we shall pursue. A *deterministic expectation function*[1] is a map F that assigns to each $x \in X$ an ordered list $\left(F^1(x), \ldots, F^{k(x)}(x)\right)$ such that $F^l(x) = \left(f^l(x), S^l(x)\right) \in X \times 2^N$ with $S^l(x) \in E\left(x, f^l(x)\right)$ for all $l = 1, \ldots, k(x)$, $S^l(x) \neq S^{l'}(x)$ for all $l \neq l'$, $f^l(x) \neq x$ for all $l \neq k(x)$, and $S^{k(x)}(x) = \emptyset$. The idea of a deterministic expectation function is that at any state x, there are some coalitions $S^l(x)$ who would replace state x by state $f^l(x)$. However, only $S^1(x)$ is allowed to actually do so. The remaining pairs in the list allow coalitions to make "rational" decisions: each coalition $S^l(x)$ knows that if they do not move to $f^l(x)$, then coalition $S^{l+1}(x)$ will move to $f^{l+1}(x)$. We assume that all players have the same expectation, represented by F, about how the abstract game unfolds, i.e., they are perfectly farsighted and have common expectations.[2] In particular, they can compare the consequences of any potential move to the consequences of not moving.

What is new in this paper is that we allow expectations to be non-deterministic. That is, each coalition might not move to a fixed new state, but can use a random device in order to decide where to move. A *probabilistic expectation function* is a map Φ that assigns to each $x \in X$ an ordered list $\left(\Phi^l(x)\right)_{l=1}^{2^{|N|}}$ such that $\Phi^l(x) = \left(\phi^l(x), S^l(x)\right) \in \Delta(X) \times 2^N$ with:

(i) $S^l(x) \neq S^{l'}(x)$ for all $l \neq l'$,

[1] Karos and Robles (2021) call this an *extended expectation function* to distinguish it from the expectation function in Dutta and Vohra (2017). The latter will play no role in this paper, so we only distinguish between deterministic and probabilistic expectation functions.

[2] For a model of heterogeneous expectations in abstract games refer to Bloch and van den Nouweland (2020).

(ii) for all $l = 1, \ldots, 2^{|N|}$, $S^l(x) \in E(x, y)$ for all $y \in \text{Supp}\left(\phi^l(x)\right)$,

(iii) $S^{2^{|N|}}(x) = \emptyset$.

Observe that this list contains all coalitions, the empty set being the last one, but, by the first condition, no coalition appears twice. The second condition ensures that any coalition S^l can only move to those states y with positive probability for which it is effective. We write $\phi^l(y \mid x)$ for the probability with which coalition $S^l(x)$ implements a move from x to y. By the last condition and since the empty set is never effective for a move out of any x, it holds that $\phi^{2^{|N|}}(x \mid x) = 1$, i.e., $\phi^{2^{|N|}} = \delta_x$.

2.3 Expected Payoffs

Let Φ be a probabilistic expectation function. Then at each $x \in X$, the probability that coalition $S^l(x)$ will implement a move to $y \in X \setminus \{x\}$ is given by $\phi^l(y \mid x) \prod_{h<l} \phi^h(x \mid x)$. Thus, the probability of a move from x to y by any coalition, i.e., the *transition probability* from x to y, is

$$p(y \mid x) = \begin{cases} \sum_{l=1}^{2^{|N|}-1} \phi^l(y \mid x) \prod_{h<l} \phi^h(x \mid x) & \text{if } y \neq x, \\ \prod_{l=1}^{2^{|N|}-1} \phi^l(x \mid x) & \text{if } y = x. \end{cases} \tag{1}$$

Let $P \in [0,1]^{X \times X}$ be the matrix with entries $P_{x,y} = p(y \mid x)$.

Lemma 2.1. *The matrix P is row-stochastic, i.e., $P_{x,y} \geq 0$ for all $x, y \in X$ and $\sum_{y \in X} P_{x,y} = 1$ for all $x \in X$.*

The proof of Lemma 2.1, as any other proof of this paper, can be found in the appendix. As P is row-stochastic, it is the transition matrix of a Markov process with state space X. Such a Markov process is called *irreducible* if for any two states x, y there is $n \in \mathbb{N}$ such that $(P^n)_{x,y} > 0$, i.e., if the probability of a transition from x to y after n steps is strictly positive. The following proposition comprises well-known results about irreducible Markov processes with finite state space that we will need later. We do not provide a proof but refer the reader to the standard literature, e.g. Stokey and Lucas (1999).

Proposition 2.2. *Let P be the transition matrix of an irreducible Markov process with finite state space X. Then there is a unique probability distribution $\mu \in \Delta(X)$*

such that for all $x \in X$

$$\mu(x) = \lim_{n\to\infty} \sum_{m=1}^{n} (P^m)_{y,x}\, e_y \tag{2}$$

for all $y \in X$*, where* e_y *is the unit vector with entry* 1 *in the* y*-th coordinate. In particular,* $\mu(x) > 0$ *for all* $x \in X$*, and* μ *satisfies* $(\mu(x))^T_{x\in X} = (\mu(x))^T_{x\in X} P$*, i.e.,* μ *is the unique (left) eigenvector of* P *to eigenvalue 1.*

Observe that μ does not depend on the choice of y on the right hand side of (2). The distribution μ is referred to as the *stationary distribution* of the Markov process. It determines for every $x \in X$ the (average) share of time that the process will spend in x. In particular, this amount is independent of the state y in which the process starts. To keep notation simple, we shall write $\mu^T P$ for the matrix product $(\mu(x))^T_{x\in X} P$.

If the expectation function Φ is such that the corresponding Markov process is irreducible, we denote the corresponding stationary distribution by μ_Φ. In this case, using the interpretation of $\mu_\Phi(x)$ as the amount of time that the process spends in x, player i will obtain the average payoff

$$u_i(\Phi) = \sum_{x\in X} \mu_\Phi(x) U_i(x). \tag{3}$$

Before we move on, we should note that the expectation functions that are investigated in the remainder of the paper will induce irreducible Markov processes; in particular, the average payoff in (3) is well-defined.

3 Hedonic Games

A *hedonic game* is a map v that maps each nonempty coalition S to some $v(S) \in \mathbb{R}^S$. That is, a hedonic game is a cooperative game such that each player's payoff in each coalition is predetermined: there is no negotiation over payoffs within coalitions whatsoever.[3] For any hedonic game v we define the map $V : \Pi \to \mathbb{R}^N$ by $V_i(\pi) = v_i(\pi(i))$. That is, $V(\pi) \in \mathbb{R}^N$ is the payoff vector for N if partition π forms.

In a hedonic game coalitions can freely form and dissolve. Thus, taking Π as the set of states, we can translate a hedonic game v into an abstract game (N, Π, E, V) with

[3]Thus, a hedonic game is an NTU-game in which no transfers among players is possible.

a suitably chosen effectivity correspondence E that represents coalitions' abilities. In particular, we would expect E to satisfy

H1 If $S \in (\pi, \pi')$, then $S \in \pi'$.

Observe that we do not allow the members of S to jointly form a partition of S: if they collaborate, they must form a coalition. If players were myopic, we would be done here, as the behavior of the remaining players were irrelevant for S's decision. But, as Ray and Vohra (2015) point out, the decision of farsighted players in S very much might depends on what is happening in $N \setminus S$, as their behavior upon the forming of S might influence future deviations. To avoid unintuitive results, they propose two conditions[4]:

H2 If $S \in E(\pi, \pi')$, $T \in \pi$, and $S \cap T = \emptyset$, then $T \in \pi'$.

H3 For every $\pi \in \Pi$ and $S \in P(N)$ there is π' with $S \in \pi'$ and $S \in E(\pi, \pi')$.

H2 requires that a coalition S that deviates from partition π has no influence over coalitions that have not been affected by its deviation. **H3** requires that from each partition π each coalition S that is not a member of π can deviate. Both conditions are highly appropriate in the context of hedonic games: they endow coalitions with the power to form at any state, yet they ensure that no coalition has the power to affect the behavior of others when moving.[5] A last observation worth making is that **H1** and **H2** imply $\pi = \pi'$ whenever $S \in \pi$ and $S \in E(\pi, \pi')$.

H1 – **H3** still allow for quite a range of partitions π' that a coalition S might move to from π, as nothing has been said about those players who where "left behind" by S. Define for any partition π and any coalition S the set $\pi(S)$ by $\pi(S) = \bigcup_{i \in S} \pi(i)$, which is the set of all players whose coalitions are affected by a deviation of S. There is no reason to presume S have power about the behavior of $\pi(S) \setminus S$. Yet, we shall assume that there is a (common) expectation about their behavior. A *residual map* is a map τ, which maps each pair (π, S) on a partition $\tau(\pi, S)$ of the set $\pi(S)$ with $S \in \tau(\pi, S)$. For $i \in \pi(S)$ we write $\tau(i \mid \pi, S)$ for the unique element of $\tau(\pi, S)$ that contains i.

[4] Ray and Vohra (2015) formulate their conditions for general NTU games; we provide here the adaption to hedonic games.

[5] This is not to say that there are no later moves that such groups might want to undertake, or that such moves are not being expected. Such moves, however, are the decisions of the moving groups at the new partition rather than a decision of the deviating coalition at the old state.

H4 There is a residual map τ such that if $S \in E(\pi, \pi')$, then $\pi'(i) = \tau(i \mid \pi, S)$ for all $i \in \pi(S)$.

H4 ensures that the behavior of $\pi(S)$ cannot be chosen by S, yet is uniquely determined and commonly known. Thus, **H2** and **H4** simply ensure that all coalitions have a common expectation about how the game will unfold after any move.

We shall not impose any conditions on the residual map; for the remainder its existence is sufficient. Yet, there are several instance of τ that have been investigated in the literature before. For instance, Hart and Kurz (1983) consider a model where coalitions who are left behind either split up or remain as they are.

Example 3.1. For any pair (π, S) of a partition and a coalition let $\gamma(\pi, S) = \left\{S, \{i\}_{i \in \pi(S) \setminus S}\right\}$. The unique effectivity correspondence that satisfies **H1**–**H4** with $\tau = \gamma$ is given by

$$E^{\gamma}(\pi, \pi') = \left\{ S \in \pi' : \begin{array}{l} T \in \pi' \text{ for all } T \in \pi \text{ with } T \cap S = \emptyset \text{ and} \\ \{\{i\}\}_{i \in T \setminus S} \subseteq \pi' \text{ for all } T \in \pi \text{ with } T \cap S \neq \emptyset \end{array} \right\}.$$

For any pair (π, S) of a partition and a coalition let $\delta(\pi, S) = \left\{S, \{\pi(i) \setminus S\}_{i \in \pi(S) \setminus S}\right\} \setminus \{\emptyset\}$. The unique effectivity correspondence that satisfies **H1**–**H4** with $\tau = \delta$ is given by

$$E^{\delta}(\pi, \pi') = \{S \in \pi' : T \setminus S \in \pi \text{ for all } T \in \pi \text{ with } T \setminus S \neq \emptyset\}.$$

Observe that in both cases for each partition π and each coalition S there is a unique partition π' with $S \in E(\pi, \pi')$. $\square$

As **H2** and **H4** together uniquely determine the behavior of $N \setminus S$ for any S and π, we obtain the following Lemma.

Lemma 3.2. *Let v be a hedonic game. Then there is a unique effectivity correspondence E that satisfies **H1**–**H4**. Moreover, for each partition π and each coalition $S \subseteq N$ there is a unique partition π' with $S \in E(\pi, \pi')$.*

Recall that we had defined players' payoffs from a probabilistic expectation function by means of the stationary distribution of a Markov process. This utility is well-defined only if the stationary distribution exists and is unique, i.e., if the process is irreducible. Hence, before we go any further and apply the ideas from Section 2, we need to understand what sequences of moves are possible, or, put differently, from what initial state to what terminal state a sequence of deviations can lead. For the above construction the next lemma provides the answer.

Lemma 3.3. *Let v be a hedonic game and let E be an effectivity correspondence that satisfies **H1**–**H3**. Then for any two partitions $\underline{\pi}$, $\overline{\pi}$, there are an integer m, partitions $\pi^1, \ldots, \pi^m$, and coalitions $S^1, \ldots, S^{m+1}$ such that $S^1 \in E(\underline{\pi}, \pi^1)$, $S^l \in E\left(\pi^{l-1}, \pi^l\right)$ for $l = 2, \ldots, m$, and $S^{m+1} \in E(\pi^m, \overline{\pi})$.*

So, the effectivity correspondence that is associated with some hedonic game does not have any two states that cannot be linked by a sequence of deviations. This is good news as it implies that probabilistic expectation functions with strictly positive probabilities wherever possible induce irreducible Markov processes.

4 Coalition Formation

4.1 Coalition Behavior

We follow Kimya (2020) and define a *pure coalition behavior* of coalition S as a map $b_S : X \to X$ with $S \in E(x, b_S(x))$ for all $x \in X$.[6] Denote the set of S's pure behaviors by B_S. A *(mixed) coalition behavior* of coalition S is a map $\beta_S : X \to \Delta(X)$ with $S \in E(x, y)$ for all $x \in X$ and all $y \in \text{Supp}(\beta_S(x))$.[7] We say that a coalition behavior is *completely mixed* if

$$\text{Supp}(\beta_S(x)) = \{y \in X : S \in E(x, y)\}.$$

With a slight abuse of notation we write $\Delta(B_S)$ for the set of all coalition behaviors of S. A *behavior profile* is a vector $(\beta_S)_{S \subseteq N} \in \prod_{S \subseteq N} \Delta(B_S)$ of behaviors.

Recall from Lemma 3.2 that an effectivity correspondence that satisfies **H1**–**H4** prescribes for each partition π and each nonempty coalition S with $S \notin \pi$ a unique

[6] We presume stationarity here: a coalition's decision at state x only depends on x and not on how or when x was reached.

[7] It is worth mentioning that our definition of mixed behaviors is more closely related to *behavior* strategies than to *mixed* strategies. This is mainly for the ease of exposition as the abstract game does not contain any imperfect information, and, hence, mixed and behavioral strategies are equivalent.

partition π' to which S can deviate. Thus, we obtain the following corollary, whose proof is omitted.

Corollary 4.1. *Let v be a hedonic game and let E be an effectivity correspondence that satisfies **H1**–**H4**. Then for any $\pi \in \Pi$, any $\emptyset \neq S \subseteq N$ with $S \notin \pi$, and any pure coalition behavior b_S it holds that $b_S(\pi) \in \{\pi, \pi'\}$, where π' is the unique partition in (7).*

By Corollary 4.1 at any partition π a coalition S has only the choice between staying at π or, if $S \notin \pi$ moving to a unique partition π'. That is, any mixed behavior of S can mix between at most two pure choices at any π'. Formally, we have the following corollary, which does not require a proof.

Corollary 4.2. *Let v be a hedonic game, and let E be an effectivity correspondence that satisfies **H1**–**H4**. Let $\emptyset \neq S \subseteq N$, let $\pi \in \Pi$, and $\beta_S \in \Delta(B_S)$. Let $\underline{\beta}_S, \bar{\beta}_S$ be such that $\underline{\beta}_S(\pi') = \bar{\beta}_S(\pi') = \beta_S(\pi')$ for all $\pi' \neq \pi$, and $\underline{\beta}_S(\pi \mid \pi) = 1$ and $\bar{\beta}_S(\pi \mid \pi) = 0$. Let $r = \beta_S(\pi \mid \pi)$. Then $\beta_S = r\underline{\beta}_S + (1-r)\bar{\beta}_S$.*

Geometrically speaking, Corollary 4.2 reveals that the set of mixed strategies of coalition S that are fixed everywhere except at π is (part of) a line; namely the line between staying at π with probability 1, or moving with probability 1 to the unique π' that S is effective for moving to.

4.2 A Coalition Formation Game

Let $\rho = (\rho_x)_{x \in X}$ be a collection of bijections $\rho_x : \left\{1, \ldots, 2^N - 1\right\} \to P(N)$. Such bijection should be interpreted as a linear order over coalitions: $\rho_x(1)$ is the first coalition, $\rho_x(2)$ is the second and so forth. We adapt the construction of Karos and Robles (2021) for our purposes and define for such an order and a behavior profile β the expectation function Φ_β by $\Phi_\beta(x) = \left(\phi_\beta^l(x), S_\beta^l(x)\right)_{l=1,\ldots,2^N}$ with

$$
\begin{aligned}
S_\beta^l(x) &= \rho_x(l) \\
\phi_\beta^l(y \mid x) &= \beta_{S_\beta^l}(y \mid x)
\end{aligned} \tag{4}
$$

for $l = 1, \ldots, 2^N - 1$, and $\Phi_\beta^{2^{|N|}}(x) = (\delta_x, \emptyset)$.

If Φ_β is such that the resulting Markov process is irreducible, then we can assign to any strategy profile β the payoff vector

$$u_i(\beta) = u_i(\Phi_\beta). \tag{5}$$

We have already seen that for effectivity correspondences that are emerging from hedonic games, we can find paths from each state to each state. In fact, there are many paths, so that even if we delete one link, all states remain connected. Thus, the following lemma is not only very useful but equally intuitive.

Lemma 4.3. *Let v be a hedonic game, let E be an effectivity correspondence that satisfies* ***H1–H4****, and let $\rho = (\rho_\pi)_{\pi\in\Pi}$ be a collection of bijections between $\{1,\ldots,2^N-1\}$ and $P(N)$. Let β be a profile of completely mixed coalition behaviors, let $\pi \in \Pi$, and let $\emptyset \neq S \subseteq N$. If $\beta'_S(\pi') = \beta_S(\pi')$ for all $\pi' \neq \pi$, then the Markov process associated with $\Phi_{\beta'_S,\beta_{-S}}$ is irreducible.*

The important consequence of the foregoing lemma is that to any behavior profile β that is completely mixed everywhere but for one coalition at one partition, we can assign a payoff vector as in (3) and (5). To make this statement more precise, define for any $\varepsilon > 0$ the set $\Delta^\varepsilon(B_S)$ as the set of all mixed coalition behaviors of S with $\beta_S(\pi' \mid \pi) \geq \varepsilon$ for all $\pi, \pi' \in \Pi$ with $S \in E(\pi,\pi')$. Then the map $u_i : \prod_S \Delta^\varepsilon(B_S) \to \mathbb{R}$ as defined by (3) and (5) is well-defined. Although $u_i(\beta)$ is quite hard to compute as the stationary distribution of the Markov process associated with Φ_β must be found, we can show that this utility function is well-behaved.

Proposition 4.4. *Let v be a hedonic game, let E be an effectivity correspondence that satisfies* ***H1–H4****, and let $\rho = (\rho_\pi)_{\pi\in\Pi}$ be a collection of bijections between $\{1,\ldots,2^N-1\}$ and $P(N)$. For every $\varepsilon > 0$ the map $u_i : \prod_S \Delta^\varepsilon(B_S) \to \mathbb{R}$ as defined by (3) and (5) is continuous.*

4.3 Weak Best Responses

From here on suppose that the hedonic game v, the permutations ρ_π, and the residual map τ that defines the effectivity correspondence E are fixed. The associated ε-*coalition formation game* is the tuple $(N, \prod_S \Delta^\varepsilon(B_S), E, V)$.

Let S be a nonempty coalition, let $\beta_{-S} = (\beta_T)_{T\neq S}$ be a profile of completely mixed behaviors for all coalitions except S, and let β_S, β'_S be two completely mixed coalition behaviors for S. We say that β_S is a *better response (in $\Delta^\varepsilon(B_S)$)* than β'_S against β_{-S} at π if $\beta_S(\pi') = \beta'_S(\pi')$ for all $\pi' \neq \pi$ and

$$u_i(\beta_S, \beta_{-S}) > u_i(\beta'_S, \beta_{-S})$$

for all $i \in S$. Essentially, β_S is a one-shot deviation from β'_S in that it differs from β'_S only at state π.

Let $\beta \in \prod_{S\subseteq N} \Delta^\varepsilon(B_S)$. We say that β^*_S is a *weak best response (in $\Delta^\varepsilon(B_S)$)* against β at π if there is no better response than β^*_S against β_{-S} at π (in $\Delta^\varepsilon(B_S)$) and $\beta^*_S(\pi') = \beta_S(\pi')$ for all $\pi' \neq \pi$. Moreover, β^*_S is a *weak best response* against β_{-S} if it is a weak best response against β_{-S} at all π. Thus, a weak best response might not be stable with respect to all profitable deviations, but it is stable with respect to all one-shot deviations. In Section 6 we shall provide an example of a weak best response that is not stable with respect to general deviations. For fixed $\varepsilon > 0$ we denote the set of S's weak best responses against β_{-S} at π by

$$R^\varepsilon_{S,\pi}(\beta) = \{\beta^*_S \in \Delta^\varepsilon(B_S) \mid \beta^*_S \text{ is a weak best response against } \beta \text{ in } \Delta^\varepsilon(B_S) \text{ at } \pi\}.$$

Proposition 4.5. *Let v be a hedonic game, and let E be an effectivity correspondence that satisfies **H1**–**H4**, let $\rho = (\rho_\pi)_{\pi\in\Pi}$ be a collection of bijections between $\{1,\ldots,2^N-1\}$ and $P(N)$. For each nonempty $S \subseteq N$, each $\beta \in \prod_T \Delta^\varepsilon(B_T)$, and each $\pi \in \Pi$ the set $R^\varepsilon_{S,\pi}(\beta)$ is nonempty and compact.*

5 Weak ε-Equilibrium

Consider the ε-coalition formation game $\left(N, \prod_S \Delta^\varepsilon(B_S), (u_i)_{i\in N}\right)$. A *weak ε-equilibrium* is a mixed behavior profile β such that for each nonempty coalition $S \subseteq N$ and each partition π it holds that β_S is a best response against β at π. That is, $\beta_S \in R^\varepsilon_{S,\pi}(\beta)$ for all $\emptyset \neq S \subseteq N$ and all $\pi \in \Pi$. We call such a profile "weak" equilibrium as it is only stable with respect to one-shot deviations, but not with respect to arbitrary deviations. In order to prove that for each $\varepsilon > 0$ a weak ε-equilibrium exists, we have to show two things: first, that the set $R^\varepsilon_{S,\pi}(\beta)$ is convex for all $(S,\pi) \in P(S) \times \Pi$ and all $\beta \in \prod_T \Delta^\varepsilon(B_T)$; second, that the correspondence $\beta \mapsto \prod_{(S,\pi^*)} R^\varepsilon_{S,\pi^*}(\beta)$ is upper hemicontinuous. By Proposition 4.5 we can then apply Kakutani's fixed point theorem to prove the existence of a fixed point of this correspondence, which, by definition, is a weak ε-equilibrium.

5.1 Convexity of the Set of Best Responses

Most coalition formation problems with similar structure are accompanied by the intrinsic difficulty that the best responses as defined above do not form a convex set. One problem is that two different best replies will lead to two different Markov processes, say with transition matrices P and Q, which in turn have stationary distributions λ and μ. While a convex combination of the two best replies will lead to a Markov process with a transition matrix that is a convex combination of P and Q, there is little that can be said about the stationary distribution of this process. In particular, it is not clear whether the emerging payoff vector will be a convex combination of the first two payoff vectors. Our first main result is that we can say something about convex combinations of two Markov processes whose transistion matrices are identical everywhere but in one row.

Theorem 5.1. *Let X be a finite set, and let $P, Q \in [0,1]^{X \times X}$ be transition matrices of irreducible Markov processes over X, so that there is x^* with $P_{x,y} = Q_{x,y}$ for all $y \in X$ and all $x \neq x^*$. Let λ and μ be the (unique) stationary distributions of P and Q, respectively. Let $r \in [0,1]$ and define*

$$t = \frac{r\mu\left(x^*\right)}{r\mu\left(x^*\right) + \left(1-r\right)\lambda\left(x^*\right)} \tag{6}$$

Then $rP + \left(1-r\right)Q$ is the transition matrix of an irreducible Markov process, and $\nu = t\lambda + \left(1-t\right)\mu$ is the unique stationary distribution of this process.

Consider a (completely mixed) strategy profile β, and fix a partition π and a coalition $\emptyset \neq S \subseteq N$. Then, for any two strategies β_S^1 and β_S^2 that coincide with β_S everywhere but in π the transition matrices of the corresponding Markov processes differ only in row π. That is, they satisfy the condition of Theorem 5.1. Formally, we obtain the following result.

Theorem 5.2. *Let v be a hedonic game, and let E be an effectivity correspondence that satisfies **H1**–**H4**, let $\rho = \left(\rho_\pi\right)_{\pi\in\Pi}$ be a collection of bijections between $\left\{1, \ldots, 2^N - 1\right\}$ and $P(N)$, and let $\varepsilon > 0$. Let $S \in P(N)$, $\beta \in \prod_{T\subseteq N} \Delta^\varepsilon\left(B_T\right)$,* *and* $\pi^* \in \Pi$, *and let* $\underline{\beta}_S, \bar{\beta}_S$ *be such that* $\underline{\beta}_S(\pi) = \bar{\beta}_S(\pi) = \beta_S$ *for all* $\pi \neq \pi^*$, *and* $\underline{\beta}_S(\pi^* \mid \pi^*) = 1$ *and* $\bar{\beta}_S(\pi^* \mid \pi^*) = 0$. *Let* $r = \beta_S(\pi^* \mid \pi^*)$. *Then*

$$u_i(\beta) = tu_i\left(\underline{\beta}_S, \beta_{-S}\right) + (1-t)\, u_i\left(\bar{\beta}_S, \beta_{-S}\right)$$

for all $i \in N$, *where* t *is defined as in (6).*

This solves the issue outlined above: as long as we focus on a coalition's one-shot deviations from a fixed partition π, the payoff vector associated with a convex combination of two one-shot deviations is a convex combination of the two payoff vectors associated with either one. However, this still does not guarantee that a the mixture of two best responses is still a best response. The reason is the following: suppose coalition $\{1,2\}$ can move from state w with payoff $u(w) = (0,0)$ to three different states, x, y, z with payoffs $U(x) = (4,0)$, $u(y) = (0,4)$, and $u(z) = (3,3)$. Then any behavior is a best response as these states cannot be Pareto ranked. However, the behavior that assigns probability $\frac{1}{2}$ to both x and y is not a best response as moving to z would be a better response.

The crucial feature in this little illustration is that there are three potential states that the coalition can move to. If it can only decide between two options, then the argument does not work: either all players will prefer one of the two options, or there is a conflict of interest, which means that no probability distributions over the two is Pareto dominated. This is good news for us: as we have seen in Corollary 4.2 all mixed behaviors β_S^1 and β_S^2 that coincide everywhere except some π^* must lie on a line. Thus, we obtain the following result.

Proposition 5.3. *Let* v *be a hedonic game, let* E *be an effectivity correspondence that satisfies* ***H1***–***H4***, *let* $\rho = (\rho_\pi)_{\pi\in\Pi}$ *be a collection of bijections between* $\left\{1, \ldots, 2^N - 1\right\}$ *and* $P(N)$, *and let* $\varepsilon > 0$. *Let* $S \in P(N)$, $\beta \in \prod_T \Delta^\varepsilon(B_T)$, *and* $\pi^* \in \Pi$. *Then* $R^\varepsilon_{S,\pi^*}(\beta)$ *is convex.*

The proof of Proposition 5.3 shows actually more. Namely, for any $S \in P(N)$, $\beta \in \prod_{T\subseteq N} \Delta^\varepsilon(B_T)$, and $\pi^* \in \Pi$, the set $R^\varepsilon_{S,\pi^*}(\beta)$ must have one of three forms: either it contains only the mixed behavior with $\beta_S(\pi^* \mid \pi^*) = \varepsilon$, or it contains only the mixed behavior with $\beta_S(\pi^* \mid \pi^*) = 1 - \varepsilon$, or it contains every mixture of the two.

5.2 Existence of Weak ε-equilibria

Recall that coalition behavior profile β is a weak ε-equilibrium if $\beta \in R^{\varepsilon}_{S,\pi}(\beta)$ for all $\emptyset \neq S \subseteq N$ and all $\pi \in \Pi$. Thus, in order to prove the existence of such equilibrium, it is sufficient to show that the correspondence that maps each coalition behavior profile β to $\prod_{(S,\pi^*)} R^{\varepsilon}_{S,\pi^*}(\beta)$ has a fixed point. We have already seen that this correspondence maps each β to a nonempty, compact, convex set. In order to apply Kakutani's fixed point theorem it is, therefore, sufficient to show that it is upper hemicontinuous.[8]

Proposition 5.4. *Let v be a hedonic game, and let E be an effectivity correspondence that satisfies **H1**–**H4**, let $\rho = (\rho_\pi)_{\pi \in \Pi}$ be a collection of bijections between $\{1, \ldots, 2^N - 1\}$ and $P(N)$, and let $\varepsilon > 0$. The correspondence $R : \prod_{S \subseteq N} \Delta^{\varepsilon}(B_S) \rightrightarrows \prod_{S \subseteq N} \Delta^{\varepsilon}(B_S)$ with $\beta \mapsto \prod_{(S,\pi^*)} R^{\varepsilon}_{S,\pi^*}(\beta)$ is upper hemicontinuous.*

The properties of the correspondence R that we have proved in Propositions 4.5, 5.3, and 5.4 allow us to apply Kakutani's fixed point theorem, so that we obtain our main result.

Theorem 5.5. *Let v be a hedonic game, and let E be an effectivity correspondence that satisfies **H1**–**H4**, let $\rho = (\rho_\pi)_{\pi \in \Pi}$ be a collection of bijections between $\{1, \ldots, 2^N - 1\}$ and $P(N)$, and let $\varepsilon > 0$. The associated coalition formation game $\left(N, \prod_S \Delta^{\varepsilon}(B_S), (u_i)_{i \in N}\right)$ obtains a weak ε-equilibrium.*

6 Discussion

6.1 Better Responses versus Weak Better Responses

We have seen that the definition of weak ε-equilibria ensures stability against one-shot deviation, but not necessarily against deviations at more than one state. So, the coalition formation games that we have defined in Section 4 lack some kind of "one-shot-principle". We shall provide an example here where a coalition does not have a one-shot deviation, i.e., is playing a weak best response, but can find a better response by changing its behavior at two states.

[8] Recall that a compact-valued correspondence $R : \beta \mapsto R(\beta)$ is upper hemicontinuous if for each converging sequence $(\beta^n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} \beta^n = \beta$, and each sequence $(\gamma^n)_{n \in \mathbb{N}}$ with $\gamma^n \in R(\beta^n)$ there is a converging subsequence $(\gamma^{n_k})_{k \in \mathbb{N}}$ with $\gamma = \lim_{k \to \infty} \gamma^{n_k} \in R(\beta)$.

Let $N = \{1,2,3\}$ and v be the hedonic game given by $v(\{1\}) = 20$, $v(\{2\}) = 0$, $v(\{3\}) = 0$, $v(\{1,2\}) = (17,14)$, $v(\{1,3\}) = (17,5)$, $v(\{2,3\}) = (15,10)$ and $v(N) = (1,18,0)$. Let E be the unique effectivity correspondence that is defined by the residual map γ in Example 3.1. Define three biljections $\rho^1,\ \rho^2, \rho^3 : \{1,\cdots,7\} \to P(N)$ by

$$\begin{aligned}
(\rho^1(1),\rho^1(2),\rho^1(3),\rho^1(4),\rho^1(5),\rho^1(6),\rho^1(7)) &= (\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}),\\
(\rho^2(1),\rho^2(2),\rho^2(3),\rho^2(4),\rho^2(5),\rho^2(6),\rho^2(7)) &= (\{1,2,3\},\{2,3\},\{1,3\},\{1,2\},\{3\},\{2\},\{1\}),\\
(\rho^3(1),\rho^3(2),\rho^3(3),\rho^3(4),\rho^3(5),\rho^3(6),\rho^3(7)) &= (\{1,2\},\{1,3\},\{2,3\},\{1,2,3\},\{1\},\{2\},\{3\}).
\end{aligned}$$

Let the set of partitions by $\Pi = \{\pi_1,\pi_2,\pi_3,\pi_4,\pi_5\}$, where $\pi_1 = \{\{1\},\{2\},\{3\}\}$, $\pi_2 = \{\{1,2\},\{3\}$, $\pi_3 = \{\{2\},\{1,3\}\}$, $\pi_4 = \{\{1\},\{2,3\}\}$, and $\pi_5 = \{N\}$. Define the collection $(\rho_\pi)_{\pi\in\Pi}$ by $\rho_{\pi_1} = \rho_{\pi_3} = \rho_{\pi_5} = \rho^1$, $\rho_{\pi_2} = \rho^2$ and $\rho_{\pi_4} = \rho^3$. Let $S_0 = \{1,2\}$, which is the coalition for which we shall find a profitable deviation which is not one-shot. Define for all $T \neq S_0$ the behavior β_T by

$$\beta_T(\pi) = \begin{cases} \pi & \text{if } T \in \pi \\ \frac{19}{20}\pi + \frac{1}{20}\pi' & \text{if } T \notin \pi \text{ and } T \in E(\pi,\pi'). \end{cases}$$

Recall that this uniquely defines β_T as for each $T \in P(N)$ and each π with $T \notin \pi$ there is exactly one π' with $T \in E(\pi,\pi')$. Behavior β prescribes for any $T \neq S_0$ at any π with $T \notin \pi$ to form and deviate from π to π' with probability $\frac{1}{20}$ and to remain at π with probability $\frac{19}{20}$.

We define the behavior of S_0 by the probabilities with which they stay at each state. Surely, coalition S_0 leave π if and only if $\pi \neq \pi_2$. So, let $\varepsilon > 0$, let $p = (p_1,\ldots,p_5)$ with $p_k \in [\varepsilon, 1-\varepsilon]$ for $k = 1,3,4,5$, and $p_2 = 1$, and define behavior $\beta^p_{S_0}$ by $\beta^p_{S_0}(\pi_k \mid \pi_k) = p_k$ for $k = 1,\ldots,5$. Then $\beta^p_{S_0}$ is uniquely determined by p. The transition matrix of the Markov process associated with profile $\left(\beta^p_{S_0}, \beta_{-S_0}\right)$ is

$$P = \begin{pmatrix}
0.857375p_1 & 1-p_1 & 0.05p_1 & 0.0475p_1 & 0.045125p_1 \\
0.0835940625 & 0.7737809375 & 0.045125 & 0.0475 & 0.05 \\
0.0975 & 0.9025\,(1-p_2) & 0.81450625p_2 & 0.045125p_2 & 0.04286875p_2 \\
0.08799375p_3 & 1-p_3 & 0.05p_3 & 0.81450625p_3 & 0.0475p_3 \\
0.142625 & 0.857375\,(1-p_4) & 0.04286875p_4 & 0.0407253125p_4 & 0.7737809375p_4
\end{pmatrix}.$$

The stationary distribution of P, μ, is given by $\mu(\pi_k) = \frac{\bar{\mu}(\pi_k)}{\sum_{l=1}^{k}\bar{\mu}(\pi_l)}$, where

$$\bar{\mu}(\pi_1) = 2.493644800\ 10^{22} - 1.929915474\ 10^{22}p_2 - 1.907014861\ 10^{22}p_3$$

$$
\begin{aligned}
&- 1.779404668\ 10^{22}p_4 + 1.475886395\ 10^{22}p_2p_3 + 1.377081367\ 10^{22}p_2p_4 \\
&- 1.052956848\ 10^{22}p_2p_3p_4 + 1.360599508\ 10^{22}p_3p_4 \\
\overline{\mu}(\pi_2) = {}& 2.621440000\ 10^{23} - 2.277208106\ 10^{23}p_1 - 2.135179264\ 10^{23}p_2 \\
&- 2.135179264\ 10^{23}p_3 - 2.028420301\ 10^{23}p_4 + 1.137161257\ 10^{23}p_1p_2p_3p_4 \\
&+ 1.843589840\ 10^{23}p_1p_2 + 1.733202232\ 10^{23}p_2p_3 + 1.647343515\ 10^{23}p_2p_4 \\
&- 1.487493405\ 10^{23}p_1p_2p_3 - 1.412391146\ 10^{23}p_1p_2p_4 - 1.333534269\ 10^{23}p_2p_3p_4 \\
&- 1.411324821\ 10^{23}p_1p_3p_4 + 1.748510978\ 10^{23}p_1p_4 + 1.647089962\ 10^{23}p_3p_4 \\
&+ 1.842392792\ 10^{23}p_1p_3 \\
\overline{\mu}(\pi_3) = {}& 1.182924800\ 10^{22} - 9.029079172\ 10^{21}p_1 - 9.012404428\ 10^{21}p_3 \\
&- 8.591357328\ 10^{21}p_4 - 4.997785432\ 10^{21}p_1p_3p_4 + 6.560586496\ 10^{21}p_1p_4 \\
&+ 6.545126296\ 10^{21}p_3p_4 + 6.878677184\ 10^{21}p_1p_3 \\
\overline{\mu}(\pi_4) = {}& 1.245184000\ 10^{22} - 9.632257218\ 10^{21}p_1 - 9.608306688\ 10^{21}p_2 \\
&- 9.101201612\ 10^{21}p_4 + 7.432904694\ 10^{21}p_1p_2 + 7.023069492\ 10^{21}p_2p_4 \\
&- 5.434523952\ 10^{21}p_1p_2p_4 + 7.042336316\ 10^{21}p_1p_4 \\
\overline{\mu}(\pi_5) = {}& 1.310720000\ 10^{22} - 1.026078331\ 10^{22}p_1 - 1.016879124\ 10^{22}p_2 \\
&- 1.008443392\ 10^{22}p_3 + 7.960009508\ 10^{21}p_1p_2 + 7.823266076\ 10^{21}p_2p_3 \\
&- 6.123530024\ 10^{21}p_1p_2p_3 + 7.893891288\ 10^{21}p_1p_3.
\end{aligned}
$$

The payoffs of players 1 and 2 are

$$
\begin{aligned}
u_1\left(\beta^p_{S_0}, \beta_{-S_0}\right) &= 20(\mu_P(\pi_1) + \mu_P(\pi_4)) + 17(\mu_P(\pi_2) + \mu_P(\pi_3)) + \mu_P(\pi_5) \\
u_2\left(\beta^p_{S_0}, \beta_{-S_0}\right) &= 14\mu_P(\pi_2) + 10\mu_P(\pi_4) + 18\mu_P(\pi_5).
\end{aligned}
$$

Let now $\varepsilon = \frac{1}{20}$ and define p^* by $p^*_k = 1 - \varepsilon$ for $k = 1, 3, 4, 5$ and $p^*_2 = 1$. Then

$$
\begin{aligned}
&\frac{d}{dp^*_1} u_1\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) > 0 &\qquad& \frac{d}{dp^*_1} u_2\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) < 0 \\
&\frac{d}{dp^*_3} u_1\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) > 0 && \frac{d}{dp^*_3} u_2\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) < 0 \\
&\frac{d}{dp^*_4} u_1\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) > 0 && \frac{d}{dp^*_4} u_2\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) < 0 \\
&\frac{d}{dp^*_5} u_1\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) < 0 && \frac{d}{dp^*_5} u_2\left(\beta^{p^*}_{S_0}, \beta_{-S_0}\right) > 0.
\end{aligned}
$$

That is, any change in any p_k^* makes exactly one player better off and one player worse off. Thus, p^* induces as weak best response.

Finally, let $\varepsilon = \frac{1}{20}$ and define $\hat{p}$ by $\hat{p}_1 = \hat{p}_4 = 1 - \varepsilon$ and $\hat{p}_3 = \hat{p}_5 = \varepsilon$. Then

$$u_1\left(\beta_{S_0}^{\hat{p}}, \beta_{-S_0}\right) = 17.72703770896 > 16.2479670393 = u_1\left(\beta_{S_0}^{p^*}, \beta_{-S_0}\right)$$
$$u_2\left(\beta_{S_0}^{\hat{p}}, \beta_{-S_0}\right) = 9.19781147654 > 7.4664207782 = u_2\left(\beta_{S_0}^{p^*}, \beta_{-S_0}\right)$$

That is, by changing their behavior both at π_3 and at π_5 both members of S_0 can strictly improve their payoffs.

6.2 Conclusion

We have shown that for the class of hedonic games there is a farsighted solution which is closely related to the rational expectation functions of Karos and Robles (2021) and Dutta and Vohra (2017). This solution incorporates the expectation of arbitrarily small but positive probabilities of making mistakes on the side of coalitions. The mathematical backbone lies in Theorem 5.1 where we show that the stationary distribution of a convex combination of irreducible Markov processes whose transition matrices differ in at most one row is a convex combination of the respective stationary distributions. This observation can also be applied to other strategic games in which the Markov process depends on mixed strategy profiles and one is interested in one-shot deviations.

A Proofs

Proof of Lemma 2.1. Surely, $P_{x,y} \geq 0$ for all $x, y \in X$. Moreover,

$$\begin{aligned}\sum_{y \in X} P_{x,y} &= \sum_{y \in X} p(y \mid x) \\ &= \sum_{y \neq x} \sum_{l=1}^{2^{|N|}-1} \phi^l\left(y \mid x\right) \prod_{h<l} \phi^h\left(x \mid x\right) + \prod_{l=1}^{2^{|N|}-1} \phi^h\left(x \mid x\right) \\ &= \sum_{l=1}^{2^{|N|}-1} \prod_{h<l} \phi^h\left(x \mid x\right) \sum_{y \neq x} \phi^l\left(y \mid x\right) + \prod_{l=1}^{2^{|N|}-1} \phi^h\left(x \mid x\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{l=1}^{2^{|N|}-1} \prod_{h<l} \phi^h (x \mid x) \left(1 - \phi^l (x \mid x)\right) + \prod_{l=1}^{2^{|N|}-1} \phi^h (x \mid x) \\
&= \sum_{l=1}^{2^{|N|}-1} \left(\prod_{h=1}^{l-1} \phi^h (x \mid x) - \prod_{h=1}^{l} \phi^h (x \mid x) \right) + \prod_{l=1}^{2^{|N|}-1} \phi^h (x \mid x) \\
&= 1 - \prod_{l=1}^{2^{|N|}-1} \phi^h (x \mid x) + \prod_{l=1}^{2^{|N|}-1} \phi^h (x \mid x) \\
&= 1,
\end{aligned}
$$

as required. ■

Proof of Lemma 3.2. Let π be a partition and S be a coalition. If $S = \emptyset$ or $S \in \pi$, then $\pi' = \pi$ by **H1** and **H2**. So, let $S \neq \emptyset$ and $S \notin \pi$. Then, by **H2** and **H4**, $S \in E(\pi, \pi')$ only if

$$
\pi' = \{\tau(i \mid \pi, S)\}_{i \in \pi(S)} \cup \{T \in \pi : T \cap S = \emptyset\}. \tag{7}
$$

By **H3**, there is some π' with $S \in E(\pi, \pi')$. Thus, $S \in E(\pi, \pi')$ if and only if π' satisfies (7). This proves the second part of the lemma. E is now uniquely defined as

$$
E(\pi, \pi') = \left\{ S \in \pi' : \begin{array}{c} \pi(i) \in \pi' \text{ for all } i \in N \setminus \pi(S) = \emptyset \\ \text{and } \pi'(i) = \tau(i \mid \pi, S) \text{ for all } i \in \pi(S) \end{array} \right\}, \tag{8}
$$

which completes the proof. ■

Proof of Lemma 3.3. Let $\pi^* = \{\{i\}\}_{i \in N}$. It is sufficient to show that the claim is true for any $\underline{\pi}$ and $\overline{\pi} = \pi^*$, as well as for any $\overline{\pi}$ and $\underline{\pi} = \pi^*$. To see the first case, observe that for any $i, j \in N$ and any partition π with $\{j\} \in \pi$, there is π' with $\{i\} \in \pi'$ and $\{i\} \in E(\pi, \pi')$ by **H3**. Moreover, $\{i\} \in \pi'$ by **H1** and $\{j\} \in \pi'$ by **H2**. Thus, the successive deviation of singletons will lead from $\underline{\pi}$ to π^*. On the other hand, let $\overline{\pi} = \{P^1, \ldots, P^m\}$, and let $\pi^l = \left\{P^1, \ldots, P^l, \{\{i\}\}_{i \in \cup_{h=l+1}^m P^h}\right\}$ for all $l = 1, \ldots, m$. Then $P^1 \in E(\pi^*, \pi^1)$ and $P^l \in E\left(\pi^{l-1}, \pi^l\right)$ for $l = 1, \ldots, m$ by **H2**. As $\overline{\pi} = \pi^m$, the proof is complete. ■

Proof of Lemma 4.3. We first show that the Markov process associated with Φ_β is irreducible. To this end note that $0 < \phi_\beta^l (\pi \mid \pi) < 1$ for all $\pi \in \Pi$ and all

$l = 1, \ldots, 2^N$. That is, at each partition π and for each coalition S, there is a positive chance that all coalitions preceding S will stay at x, so that S will be able to implement its move. In particular, there is a strictly positive chance that S will actually implement its move out of π. This is, particularly, true for the singletons and the coalitions in the proof of Lemma 3.3. Thus, for any $\pi, \pi' \in \Pi$, there is a positive chance of a move from π to π', that is, $(P^m)_{\pi,\pi'} > 0$ for some $m \in \mathbb{N}$. Hence, the process associated with Φ_β is irreducible. By construction of β' the transition matrices of the processes Φ_β and $\Phi_{\beta'_S,\beta_{-S}}$ differ at most in π. Since $|N| \geq 3$ there are at every partition π' at least two coalitions that can deviate: if π is the grand coalition, every singleton can deviate; if π contains only singletons, then at least all pairs can deviate; and if π neither consists only of singletons nor the grand coalition, then the grand coalition and at least one singleton can deviate. As the order in which the deviations of singletons and the coalitions in the proof of Lemma 3.3 is irrelevant, one can always find a path from $\underline{\pi}$ to $\bar{\pi}$ that does not involve a deviation by S at π. Hence, this path will occur with possible probability, which makes $\Phi_{\beta'_S,\beta_{-S}}$ irreducible. ■

Proof of Proposition 4.4. By (3) it is sufficient to show that the map $\beta \mapsto \mu_{\Phi_\beta}$ is continuous. By Theorem 12.13 in Stokey and Lucas (1999), and since the state space Π is finite and the Markov process associated with the expectation function Φ_β is irreducible, it is sufficient to show that the map $\beta \mapsto P^{\Phi_\beta}$ is continuous, where P^{Φ_β} is the corresponding transition matrix. But this is clear, since by (1) and (4) it holds that

$$P^{\Phi_\beta}_{\pi,\pi'} = \begin{cases} \sum_{l=1}^{2^{|N|}-1} \beta_{\rho_\pi(l)}\left(\pi' \mid \pi\right) \prod_{h<l} \beta_{\rho_\pi(h)}\left(\pi \mid \pi\right) & \text{if } \pi' \neq \pi, \\ \prod_{l=1}^{2^{|N|}-1} \beta_{\rho_\pi(l)}\left(\pi \mid \pi\right) & \text{if } \pi' = \pi, \end{cases}$$

which is continuous in β for all $\pi, \pi' \in \Pi$. ■

Proof of Proposition 4.5. Let $S \in P(N)$, let $\beta \in \prod_{T \subseteq N} \Delta^\varepsilon\left(B_T\right)$, and $\pi \in \Pi$ be fixed but arbitrary. Let $i \in S$ and recall from Proposition 4.4 that u_i is continuous on $\prod_{T \subseteq N} \Delta^\varepsilon\left(B_T\right)$. For $t \in [\varepsilon, 1-\varepsilon]$ let

$$\beta_S^t\left(\pi' \mid \pi'\right) = \begin{cases} t & \text{if } \pi' = \pi \\ \beta_S\left(\pi' \mid \pi'\right) & \text{otherwise} \end{cases}$$

and recall from Corollary 4.1 that β_S^t is uniquely determined by t. As β_S^t continuously depends on t, it holds that $\hat{u}_i(t) = u_i(\beta_S^t, \beta_{-S})$ continuously depends on t; and as $[\varepsilon, 1-\varepsilon]$ is compact, $\hat{u}_i$ obtains its maximum at some t^*. By construction, there is no better response than $\beta_S^{t^*}$ against β_{-S} at π as there is no behavior that would provide a higher payoff to i. Hence, $\beta_S^{t^*} \in R_{S,\pi}^{\varepsilon}(\beta)$, and the latter is nonempty.

For compactness it is sufficient to show closeness as $\Delta^{\varepsilon}(B_S)$ is compact. For this purpose, let $(\beta_S^n)_{n\in\mathbb{N}}$ be a converging sequence in $R_{S,\pi}^{\varepsilon}(\beta)$ with limit β_S^*. Assume that $\beta_S^* \notin R_{S,\pi}^{\varepsilon}(\beta)$, i.e., assume that there is $\beta_S \in \Delta^{\varepsilon}(B_S)$ such that $\beta_S(\pi') = \beta_S^*(\pi')$ for all $\pi' \neq \pi$ and $u_i(\beta_S, \beta_{-S}) > u_i(\beta_S^*, \beta_{-S})$ for all $i \in S$. Let $\delta = \min_{i\in S} u_i(\beta_S, \beta_{-S}) - u_i(\beta_S^*, \beta_{-S}) > 0$. By the continuity of u_i there is $c > 0$ such that for all $\beta_S' \in \Delta^{\varepsilon}(B_S)$ with $\|\beta_S^* - \beta_S'\| < c$ it holds that $|u_i(\beta_S^*, \beta_{-S}) - u_i(\beta_S', \beta_{-S})| < \frac{1}{2}\delta$ for all $i \in S$. As (β_S^n) is converging, there is m such that $\|\beta_S^* - \beta_S^n\| < c$ for all $n \geq m$, so that $|u_i(\beta_S^*, \beta_{-S}) - u_i(\beta_S^n, \beta_{-S})| < \frac{1}{2}\delta$ for all $n \geq m$ and all $i \in S$. In particular, $u_i(\beta_S^n, \beta_{-S}) < u_i(\beta_S^*, \beta_{-S}) + \frac{1}{2}\delta \leq u_i(\beta_S, \beta_{-S}) - \frac{1}{2}\delta$ for all $i \in S$. But this is impossible since $\beta_S^n \in R_{S,\pi}^{\varepsilon}(\beta)$. Thus, $\beta_S^* \in R_{S,\pi}^{\varepsilon}(\beta)$. ■

Proof of Theorem 5.1. Surely, the new Markov with transition matrix $rP + (1-r)Q$ is irreducible. Thus, by Proposition 2.2 it has a unique stationary distribution, and the stationary distribution is the unique normalized left eigenvevtor to eigenvalue 1. So, it is sufficient to show that ν is a normalized left eigenvevtor to eigenvalue 1. By construction, $\sum_{x\in X} \nu(x) = 1$ and $\nu(x) > 0$ for all $x \in X$. In particular, since $r\mu(x^*) + (1-r)\lambda(x^*) = \nu(x^*) > 0$, we have

$$\begin{aligned}
&(1-r)\,t\lambda(x^*) - r(1-t)\,\mu(x^*) \\
&\qquad = \frac{(1-r)\,t\lambda(x^*) - r(1-t)\,\mu(x^*)}{r\mu(x^*) + (1-r)\,\lambda(x^*)}\left(r\mu(x^*) + (1-r)\,\lambda(x^*)\right) \\
&\qquad = \left((1-t)\,t - t(1-t)\right)\left(r\mu(x^*) + (1-r)\,\lambda(x^*)\right) \\
&\qquad = 0.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left(\nu^T\left(rP + (1-r)\,Q\right)\right)_y &= \sum_{x\in X}\left(t\lambda(x) + (1-t)\,\mu(x)\right)\left(rP_{x,y} + (1-r)\,Q_{x,y}\right) \\
&= rt\lambda(y) + (1-r)(1-t)\,\mu(y)
\end{aligned}$$

$$
\begin{aligned}
&+ (1-r)\, t \sum_{x \in X} \lambda(x) Q_{x,y} + r\,(1-t) \sum_{x \in X} \mu(x) P_{x,y} \\
&= \nu(y) + (1-r)\, t \sum_{x \in X} \lambda(x) \left(Q_{x,y} - P_{x,y}\right) \\
&\quad + r\,(1-t) \sum_{x \in X} \mu(x) \left(P_{x,y} - Q_{x,y}\right) \\
&= \nu(y) + (1-r)\, t \lambda\left(x^*\right) \left(Q_{x^*,y} - P_{x^*,y}\right) \\
&\quad - r\,(1-t)\, \mu\left(x^*\right) \left(Q_{x^*,y} - P_{x^*,y}\right) \\
&= \nu(y) + \left(Q_{x^*,y} - P_{x^*,y}\right) \left((1-r)\, t \lambda\left(x^*\right) - r\,(1-t)\, \mu\left(x^*\right)\right) \\
&= \nu(y),
\end{aligned}
$$

which proves that ν is the stationary distribution of $\alpha P + (1-\alpha)\, Q$. ■

Proof of Theorem 5.2. Let $\underline{P} = P^{\Phi\left(\underline{\beta}_S, \beta_{-S}\right)}$, let $\overline{P} = P^{\Phi\left(\bar{\beta}_S, \beta_{-S}\right)}$, and observe that the corresponding Markov processes are irreducible by Lemma 4.3. Moreover, $\underline{P}_{\pi,\pi'} = \overline{P}_{\pi,\pi'}$ for all $\pi' \in \Pi$ and all $\pi \neq \pi^*$. By construction we have $\beta_S = r \underline{\beta}_S + (1-r)\, \bar{\beta}_S$, and thus, for the transition probability of the Markov process that emerges from Φ_β we find for all $\pi' \in \Pi$ and all $\pi \neq \pi^*$

$$
P^{\Phi_\beta}_{\pi,\pi'} = \underline{P}_{\pi,\pi'} = r \underline{P}_{\pi,\pi'} + (1-r)\, \overline{P}_{\pi,\pi'}.
$$

Let l^* be such that $\rho_{\pi^*}\left(l^*\right) = S$. Then

$$
\begin{aligned}
P^{\Phi_\beta}_{\pi^*,\pi^*} &= \left(r \underline{\beta}_S \left(\pi^* \mid \pi^*\right) + (1-r)\, \bar{\beta}_S \left(\pi^*, \pi^*\right)\right) \prod_{l \neq l^*}^{2^{|N|-1}} \beta_{\rho(l)} \left(\pi^* \mid \pi^*\right) \\
&= r \underline{P}_{\pi^*,\pi^*} + (1-r)\, \overline{P}_{\pi^*,\pi^*}.
\end{aligned}
$$

Moreover, for $\pi \neq \pi^*$ we have

$$
\begin{aligned}
P^{\Phi_\beta}_{\pi^*,\pi} = &\sum_{l<l^*} \beta_{\rho(l)} \left(\pi \mid \pi^*\right) \prod_{h<l} \beta_{\rho(h)} \left(\pi^* \mid \pi^*\right) \\
&+ \left(r \underline{\beta}_{\rho(l^*)} \left(\pi \mid \pi^*\right) + (1-r)\, \bar{\beta}_{\rho(l^*)} \left(\pi \mid \pi^*\right)\right) \prod_{h<l^*} \beta_{\rho(l)} \left(\pi^* \mid \pi^*\right) \\
&+ \sum_{l>l^*} \beta_{\rho(l)} \left(\pi \mid \pi^*\right) \prod_{h<l, h \neq l^*} \beta_{\rho(h)} \left(\pi^* \mid \pi^*\right)
\end{aligned}
$$

$$
\begin{aligned}
&\qquad\qquad \cdot \left(r \underline{\beta}_{\rho(l^*)} (\pi \mid \pi^*) + (1-r) \bar{\beta}_{\rho(l^*)} (\pi \mid \pi^*) \right) \\
&= r \Bigg(\sum_{l<l^*} \beta_{\rho(l)} (\pi \mid \pi^*) \prod_{h<l} \beta_{\rho(h)} (\pi^* \mid \pi^*) \\
&\quad + \underline{\beta}_{\rho(l^*)} (\pi \mid \pi^*) \prod_{h<l^*} \beta_{\rho(l)} (\pi^* \mid \pi^*) \\
&\quad + \sum_{l>l^*} \beta_{\rho(l)} (\pi \mid \pi^*) \underline{\beta}_{\rho(l^*)} (\pi \mid \pi^*) \prod_{h<l, h \neq l^*} \beta_{\rho(h)} (\pi^* \mid \pi^*) \Bigg) \\
&+ (1-r) \Bigg(\sum_{l<l^*} \beta_{\rho(l)} (\pi \mid \pi^*) \prod_{h<l} \beta_{\rho(h)} (\pi^* \mid \pi^*) \\
&\quad + \bar{\beta}_{\rho(l^*)} (\pi \mid \pi^*) \prod_{h<l^*} \beta_{\rho(l)} (\pi^* \mid \pi^*) \\
&\quad + \sum_{l>l^*} \beta_{\rho(l)} (\pi \mid \pi^*) \bar{\beta}_{\rho(l^*)} (\pi \mid \pi^*) \prod_{h<l, h \neq l^*} \beta_{\rho(h)} (\pi^* \mid \pi^*) \Bigg) \\
&= r \underline{P}_{\pi^*, \pi} + (1-r) \overline{P}_{\pi^*, \pi}.
\end{aligned}
$$

Hence, we find that $P^{\Phi_\beta} = r\underline{P} + (1-r)\overline{P}$. Thus, by Theorem 5.1, the Markov process associated with behavior profile $\beta = \left(r\underline{\beta}_S + (1-r) \bar{\beta}_S, \beta_{-S} \right)$ has the unique stationary distribution $\nu = t\lambda + (1-t)\mu$, where λ and μ are the stationary distributions of $\underline{P}$ and $\overline{P}$, respectively. Thus, by Equations (3) and (5), we obtain

$$
\begin{aligned}
u_i(\beta) &= u_i \left(r\underline{\beta}_S + (1-r)\bar{\beta}_S, \beta_{-S} \right) \\
&= \sum_{\pi \in \Pi} \nu(\pi) U_i(\pi) = \sum_{\pi \in \Pi} t\lambda(\pi) V_i(\pi) + (1-t)\mu(\pi) V_i(\pi) \\
&= t u_i \left(\underline{\beta}_S, \beta_{-S} \right) + (1-t) u_i \left(\bar{\beta}_S, \beta_{-S} \right)
\end{aligned}
$$

for all $i \in N$. ■

Proof of Proposition 5.3. Let $\underline{\beta}_S$ and $\bar{\beta}_S$ be defined as in Theorem 5.2. If $R^{\varepsilon}_{S,\pi^*}(\beta)$ contains only one element, there is nothing left to show; so, let $\beta^1_S, \beta^2_S \in R^{\varepsilon}_{S,\pi^*}(\beta)$ with $\beta^1_S \neq \beta^2_S$. Let $r_1, r_2 \in (\varepsilon, 1-\varepsilon)$ be such that $\beta^l_S = r_l \underline{\beta}_S + (1-r_l)\bar{\beta}_S$ for $l = 1, 2$. By Theorem 5.2 it holds that

$$
u_i \left(\beta^l_S, \beta_{-S} \right) = t_l u_i \left(\underline{\beta}_S, \beta_{-S} \right) + (1-t_l) u_i \left(\bar{\beta}_S, \beta_{-S} \right)
$$

for $l = 1, 2$, where $t_l \in (0,1)$ is defined as in (6) for $r = r_l$. Without loss of generality assume that $r_1 > r_2$ and observe that this is equivalent to $t_1 > t_2$. Since both β_S^1 and β_S^2 are best responses against β at π^*, there are $i, j \in S$ (possibly with $i = j$), such that $u_i\left(\beta_S^1, \beta_{-S}\right) \geq u_i\left(\beta_S^2, \beta_{-S}\right)$ and $u_j\left(\beta_S^2, \beta_{-S}\right) \geq u_j\left(\beta_S^1, \beta_{-S}\right)$. In particular, since $t_1 > t_2$, this means

$$u_i\left(\underline{\beta}_S, \beta_{-S}\right) \geq u_i\left(\bar{\beta}_S, \beta_{-S}\right) \tag{9}$$

$$u_j\left(\underline{\beta}_S, \beta_{-S}\right) \leq u_j\left(\bar{\beta}_S, \beta_{-S}\right). \tag{10}$$

Let now $q \in (0,1)$ and $\hat{\beta}_S = q\beta_S^1 + (1-q)\,\beta_S^2$. Then

$$\hat{\beta}_S = \hat{r}\underline{\beta}_S + (1-\hat{r})\,\bar{\beta}_S.$$

where $\hat{r} = qr_1 + (1-q)r_2$. Assume that $\hat{\beta}_S \notin R^{\varepsilon}_{S,\pi^*}(\beta)$. Then there is $\beta_S^* \in \Delta^{\varepsilon}(B_S)$ such that $\beta_S^*(\pi) = \beta_S(\pi)$ for all $\pi \neq \pi^*$ and $u_k\left(\beta_S^*, \beta_{-S}\right) > u_k\left(\hat{\beta}_S, \beta_{-S}\right)$ for all $k \in S$. By Corollary 4.2 there is $r^* \in \mathbb{R}$ such that $\beta_S^* = r^*\underline{\beta}_S + (1-r^*)\,\bar{\beta}_S$. Define $\hat{t}$ and t^* as in (6) for $\hat{r}$ and r^*, respectively, If $t^* \geq \hat{t}$, then, by Theorem 5.2 and (10),

$$\begin{aligned} u_j\left(\beta_S^*, \beta_{-S}\right) &> u_j\left(\hat{\beta}_S, \beta_{-S}\right) = \hat{t}u_j\left(\underline{\beta}_S, \beta_{-S}\right) + \left(1-\hat{t}\right)u_j\left(\bar{\beta}_S, \beta_{-S}\right) \\ &\geq t^*u_j\left(\underline{\beta}_S, \beta_{-S}\right) + (1-t^*)\,u_j\left(\bar{\beta}_S, \beta_{-S}\right) = u_j\left(\beta_S^*, \beta_{-S}\right), \end{aligned}$$

which is impossible; and if $t^* \leq \hat{t}$, then, similarly with (9),

$$\begin{aligned} u_i\left(\beta_S^*, \beta_{-S}\right) &> u_i\left(\hat{\beta}_S, \beta_{-S}\right) = \hat{t}u_i\left(\underline{\beta}_S, \beta_{-S}\right) + \left(1-\hat{t}\right)u_i\left(\bar{\beta}_S, \beta_{-S}\right) \\ &\geq t^*u_i\left(\underline{\beta}_S, \beta_{-S}\right) + (1-t^*)\,u_i\left(\bar{\beta}_S, \beta_{-S}\right) = u_i\left(\beta_S^*, \beta_{-S}\right), \end{aligned}$$

which is impossible as well. ■

Proof of Proposition 5.4. Let $(\beta^n)_{n\in\mathbb{N}}$ be a converging sequence of mixed behavior profiles $\beta^n \in \prod_T \Delta^{\varepsilon}(B_T)$ with $\lim_{n\to\infty}\beta^n = \beta$, and let $(\gamma^n)_{n\in\mathbb{N}}$ be a sequence with $\gamma^n \in R(\beta^n)$ for all $n \in \mathbb{N}$. As $R(\beta^n) \subseteq \prod_T \Delta^{\varepsilon}(B_T)$ and the latter is compact, there is a converging subsequence $(\gamma^{n_k})_{k\in\mathbb{N}}$ with $\gamma = \lim_{k\to\infty}\gamma^{n_k} \in \prod_T \Delta^{\varepsilon}(B_T)$. Assume that $\gamma \notin R(\beta)$. Then there are $S \in P(N)$ and $\pi^* \in \Pi$ such that $\gamma_S \notin R^{\varepsilon}_{S,\pi^*}(\beta)$. Thus, there is $\alpha_S \in \Delta^{\varepsilon}(B_S)$, such that $\alpha_S(\pi) = \gamma_S(\pi)$ for all $\pi \neq \pi^*$ and

$u_i(\alpha_S, \beta_{-S}) > u_i(\gamma_S, \beta_{-S})$ for all $i \in S$. Let $\left(\alpha_S^k\right)_{k\in\mathbb{N}}$ be a sequence in $\Delta^{\varepsilon}(B_S)$ such that $\lim_{k\to\infty} \alpha_S^k = \alpha_S$. Moreover, let $\delta = \min_{i\in S} u_i(\alpha_S, \beta_{-S}) - u_i(\gamma_S, \beta_{-S}) > 0$. By the continuity of u_i there is K^1 such that $\left|u_i\left(\gamma_S^{n_k}, \beta_{-S}^{n_k}\right) - u_i(\gamma_S, \beta_{-S})\right| < \frac{1}{2}\delta$ for all $k \geq K^1$ and all $i \in S$. Similarly, there is K^2 such that $\left|u_i\left(\alpha_S^k, \beta_{-S}\right) - u_i\left(\alpha_S, \beta_{-S}^{n_k}\right)\right| < \frac{1}{2}\delta$ for all $k \geq K^2$ and all $i \in S$. Thus,

$$u_i\left(\alpha_S^k, \beta_{-S}^{n_k}\right) > u_i(\alpha_S, \beta_{-S}) - \frac{1}{2}\delta \geq u_i(\gamma_S, \beta_{-S}) + \frac{1}{2}\delta > u_i\left(\gamma_S^{n_k}, \beta_{-S}^{n_k}\right)$$

for all $k \geq \max\{K^1, K^2\}$ and all $i \in S$. But this is a contradiction as $\gamma^{n_k} \in R_{S,\pi^*}(\beta^{n_k})$ by construction. Hence, $\gamma \in R(\beta)$, which proves upper hemicontinuity. ■

References

Banerjee, S., H., K., Sönmez, T., 2001. Core in a simple coalition formation game. Social Choice and Welfare 19, 135–153.

Bloch, F., van den Nouweland, A., 2020. Farsighted stability with heterogeneous expectations. Games and Economic Behavior 121, 32–54.

Bogomolnaia, A., Jackson, M., 2002. The Stability of Hedonic Coalition Structurs. Games and Economic Behavior 38, 201–230.

Bondareva, O.N., 1963. Some applications of linear programming methods to the theory of cooperative games. Problemy Kybernetiki 10, 119–139.

Chwe, M.S.Y., 1994. Farsighted coalitional stability. Journal of Economic Theory 63, 299–325.

Diamantoudi, E., Xue, L., 2003. Farsighted stability in hedonic games. Social Choice and Welfare 21, 39–61.

Drèze, J., Greenberg, J., 1980. Hedonic coalitions: Optimalilty and stability. Econometrica 48, 987–1003.

Dutta, B., Vohra, R., 2017. Rational expectations and farsighted stability. Theoretical Economics 12, 1191–1227.

Harsanyi, J.C., 1974. An equilibrium-point interpretation of stable sets and a proposed alternative definition. Management Science 20, 1472–1495.

Hart, S., Kurz, M., 1983. Endogenous formation of coalitions. Econometrica: Journal of the Econometric Society , 1047–1064.

Iehlé, V., 2007. The core-partition of hedonic games. Mathematical Social Sciences 54, 176–185.

Karos, D., Robles, L., 2021. Full farsighted rationality. Games and Economic Behavior, 130, 409–424.

Kimya, M., 2020. Equilibrium coalitional behavior. Theoretical Economics 15, 669–714.

Ray, D., Vohra, R., 2015. The farsighted stable set. Econometrica 83, 977–1011.

Shapley, L., 1967. On balanced sets and the core. Naval Research Logistics Quarterly 14, 453–460.

Stokey, N.L., Lucas, R.E.J., 1999. Recursive Methods in Economic Dynamics, 5th edition. Havard University Press.