A BARGAINING APPROACH TO COORDINATION IN NETWORKS*

Miguel A. Meléndez**

WP-AD 2007-28

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Diciembre 2007

Depósito Legal: V-212-2008

Ivie working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

---

* This is a revised version of Chapter 1 of my PhD Thesis presented at Universidad de Alicante. A previous version of this paper circulated with the title "Network Formation and Coordination: Bargaining the Division of Link Costs" (IVIE WP-AD 2002-27). I am indebted to Fernando Vega-Redondo for many helpful conversations, suggestions and encouragement. I am grateful to Julio González-Díaz, Antonio J. Morales and Ascensión Andina-Díaz for valuable suggestions. I also thank Jordi Brandts, Antonio Cabrales, Angel Hernando, Andrea Galeotti, Sanjeev Goyal, Christoph Kuzmics, Ana Lozano-Vivas, Paola Manzini, Juan Mora, Efe Ok, Juana Santamaría-García and Jose A. Silva for their comments. I gratefully acknowledge the hospitality of New York University, where part of this research was carried out, and the financial support from the Spanish Ministry of Education. The usual disclaimer applies.

** Departamento de Teoría e Historia Económica, Universidad de Málaga, 29013 Málaga, Spain. E-mail: melendez@uma.es.

A BARGAINING APPROACH TO COORDINATION IN NETWORKS

Miguel A. Meléndez

ABSTRACT

This paper presents a model in which players interact via the formation of costly links and the benefits of bilateral interactions are determined by a coordination game. A novel contribution of this paper is that the fraction of the cost borne by each player involved in a bilateral link is not fixed exogenously, but results from bargaining. We analyze the model both in a static and a dynamic setting. Whereas the static game has multiplicity of equilibria, we show that only one is stochastically stable.

Keywords: Coordination; Nash bargaining solution; Risk-dominance; Stochastic stability.

JEL Classification: C72; C73; C78

1 Introduction

In many situations, in order to be able to interact, different agents need to agree on how to share some costs. Consider for example two firms that plan to engage in a cooperative agreement, namely, a laundry firm and a student's residence. The laundry firm would benefit from having some room inside the residence, since it will get new clients. The residence would also benefit, since providing laundry service might increase the occupancy in the future. But, how should the cost of adapting some space for the laundry machines be borne?

The literature on network formation studies situations in which agents can connect with each other in order to get benefits.[1] If link formation is costly, there are two kind of models regarding how these costs are met: *one-sided* models, in which the agent that proposes to form a link meets the entire cost, and *two-sided* models, in which both agents share the cost evenly.[2] In both models the fraction of the cost that each agent bears in a link is exogenous. Nevertheless, when two players have the possibility to form a link, there are several possibilities to split the cost it involves. In our model, the cost shares are decided through a bargaining process that depends on the potential benefits they can get from interaction and on what the two agents get in the case the link is not formed. In this sense, we make the cost sharing endogenous.[3]

One can think of different sources of benefits from establishing links. For example, the remarkable paper of Bala and Goyal (2000) models situations where profits come from the mere fact of communication. In this paper, we analyze a network formation model in which the potential benefits of each pair of connected agents depend on a coordination game. Other papers that study the formation of networks in contexts of coordination games are Goyal and Vega-Redondo (2005), Jackson and Watts (2002) and Droste *et al.* (2000).[4] Our coordination game has two pure Nash equilibria: one efficient and the other risk-dominant. Our main objective is to study how a dynamic process affects the likelihood of the different equilibria.

Our game has two stages. First, each agent chooses an action. Second, after observing the action profile, agents bargain (on a bilateral basis) on the formation of links and the corresponding cost sharing; we assume that the bargaining outcome is determined by the Nash solution. We show that, in the static model, both the efficient and the risk-dominant action profiles can be supported in equilibrium. Moreover, in

[1] See Jackson (2005) for a survey of the literature of network formation.

[2] For example, Bala and Goyal (2000), Goyal and Vega-Redondo (2005), Bramoulle *et al.* (2004), and Galeotti *et al.* (2006) consider one-sided models. Examples of two-sided models are Watts (2001), Jackson and Watts (2002), Droste *et al.* (2000) and Jackson and Wolinsky's (1996) connections model.

[3] In the setup of communication networks, Slikker and van den Nouweland (2000) consider multilateral bargaining over the division of the benefits and the costs simultaneously, in a 3-person cooperative game.

[4] Vega-Redondo (2005), Haag and Lagunoff (2004) and Galeotti and Meléndez-Jiménez (2004) consider a prisoner's dilemma game and Bramoulle *et al.* (2004) an anti-coordination game.

equilibrium, either the complete or the empty network forms and, in every formed link, the cost is shared evenly.

The dynamic version of the model addresses the issue of equilibrium selection. We show that there is multiplicity of limit states, although the range of parameters for which there is multiplicity shrinks with respect to the static game. Then, in order to examine the robustness of each limit state, we characterize the set of stochastically stable states.[5] The main result of the paper (Theorem 2) shows that the *risk-dominance premium* of the bilateral coordination game uniquely determines a stochastically stable state. The risk-dominance premium is a particularization of the optimization premium in Battalio *et al.* (2001) for the case in which a player expects his opponent to play each action with equal probability in the bilateral game. Intuitively, the risk-dominance premium is a measure of the degree of risk-dominance of the coordination game. We show that if the risk-dominance premium exceeds the share of the cost that a player meets in equilibrium, all players choose the risk-dominant action in the unique stochastically stable state and, otherwise, the efficient action is selected. The equilibrium network in the long run is the complete network and the costs of the links are shared evenly.

We finally study how this result is altered when we relax some of the assumptions of the paper. We first allow for side payments in the link formation and then allow the payoffs of the bilateral coordination game to be negative. In both cases we obtain that the unique stochastically stable state is the efficient one, regardless of the risk-dominance premium. We then compare our results to other relevant papers that also analyze network formation in a context of coordination games, but assume a different cost structure.

The remainder of the paper is organized as follows. In Section 2, we present the static model, analyze the bargaining outcome and characterize the equilibria. In Section 3, we introduce the dynamics, characterize the limit sets and identify the stochastically stable states. In Section 4, we relax the main assumptions of the model. Finally, in Section 5, we compare our results to the most related papers.

2 The static model

Let $N := \{1, ..., n\}$ be the set of players, where $n > 2$ and let $G := \{g \subset N \times N :$ for all $i, j \in N$, $(i, i) \notin g$ and $(i, j) \in g \Longleftrightarrow (j, i) \in g\}$ be the set of undirected networks. We consider a two-stage game. In the first stage, each player $i \in N$ chooses an action $a_i \in \{\alpha, \beta\}$. Then, all players observe the action profile $a = (a_1, ..., a_n)$ and, in the second stage, players form a network $g \in G$. Let $A := \{\alpha, \beta\}^n$ be the set of action profiles. Given $a \in A$, the establishment of link (i, j) implies:

[5] The literature which analyzes stochastic stability in contexts of coordination games is extensive. For instance, see Kandori et al. (1993), Young (1993), Ellison (1993, 2000), Ely (2002), Baskhar and Vega-Redondo (2004), Goyal and Vega-Redondo (2005), Jackson and Watts (2002), Droste *et al.* (2000) and Peski (2003, 2004).

i) A gross payoff for each of the two players, which is given by the symmetric function $\pi : \{\alpha, \beta\} \times \{\alpha, \beta\} \to \mathbb{R}$, which only depends on the actions chosen in the first stage: $\pi(a_i, a_j)$, for player i, and $\pi(a_j, a_i)$, for player j. The function π is represented in Table 1 below, where $d > f$, $b > e$, $d > b$ and $b + f > d + e$. The conditions $d > f$ and $b > e$ reflect the fact that $\pi(\cdot)$ represents a 2×2 symmetric coordination game. The condition $d > b$ implies that (α, α) is its Pareto efficient equilibrium. Finally, the condition $b+f > d+e$ implies that (β, β) is the risk-dominant equilibrium, as defined by Harsanyi and Selten (1988). This equilibrium has the property that each player is choosing a best response to the other player mixing 50/50 in the coordination game. We also assume that $e > 0$, *i.e.*, all the entries of the payoff matrix are positive.[6]

	α	β
α	d	e
β	f	b

Table 1: the bilateral (gross) payoff matrix.

ii) A fixed cost $c > 0$, to be borne by players i and j. We assume $c < 2d$, since, otherwise, no link is ever formed.

Now we discuss how the link formation takes place. Given $a \in A$, each pair $i, j \in N$, $i \neq j$, bargain to share the surplus they would get by forming the link (i, j), *i.e.*, $\pi(a_i, a_j) + \pi(a_j, a_i) - c$. Let $\hat{\pi}(a_i, a_j)$ represent the net payoff that player i for being linked to player j. For most of the paper, we assume that side payments are not allowed, *i.e.*, $\hat{\pi}(a_i, a_j) \leq \pi(a_i, a_j)$.[7] If $\pi(a_i, a_j) + \pi(a_j, a_i) \leq c$, then $(i, j) \notin g$ since players i and j can not benefit from forming the link. In such a case, $\hat{\pi}(a_i, a_j) := 0$ and $\hat{\pi}(a_j, a_i) := 0$. Suppose now that $\pi(a_i, a_j) + \pi(a_j, a_i) > c$, then i and j bargain to share the surplus. In such a case, $B(a_i, a_j) := \{(x_i, x_j) \in \mathbb{R}^2 : x_i + x_j \leq \pi(a_i, a_j) + \pi(a_j, a_i) - c,\ x_i \leq \pi(a_i, a_j) \text{ and } x_j \leq \pi(a_j, a_i)\}$ defines a bargaining set and the pair $(B(a_i, a_j), (0, 0))$ represents a two-person bargaining problem, where $(0, 0)$ is the disagreement value.[8] We assume that, if players i and j reach an agreement, the link (i, j) is formed and the outcome is determined by the Nash solution, that is, $(\hat{\pi}(a_i, a_j), \hat{\pi}(a_j, a_i)) := \mathcal{NS}(B(a_i, a_j), (0, 0)) = \arg\max_{(x_i, x_j) \in B(a_i, a_j), x_i, x_j \geq 0} x_i \cdot x_j$. Otherwise, the link is not formed and the outcome is the disagreement value. When the link (i, j) is formed, the cost of player i is $c(a_i, a_j) = \pi(a_i, a_j) - \hat{\pi}(a_i, a_j)$.

Therefore, the timing of the game is as follows:

1. Players simultaneously choose actions. Then, the chosen action profile $a \in A$ is made public.

[6] This assumption is relaxed in Section 4.

[7] In section 4, we analyze how allowing for side payments affects the main result of the paper.

[8] A two-person bargaining problem is a pair (B, v) such that: i) B is a convex, closed and comprehensive subset of $\mathbb{R}^2$, ii) $v \in B$ and there exists $x \in B$ such that $x > v$ and iii) $B_v = \{x \in B : x \geq v\}$ is a compact set. B is called the bargaining set and v the disagreement value.

2. Network formation. Each $i \in N$ chooses $S^i(a) \subseteq N \backslash \{i\}$. Given $i, j \in N$, $(i,j) \in g$ if and only if $j \in S^i(a)$ and $i \in S^j(a)$.[9]

3. Payoffs are realized. Each $i \in N$ gets, for each $j \in N \backslash \{i\}$, the following net payoff: 0 if $(i,j) \notin g$, and $\mathcal{NS}(B(a_i, a_j), (0,0))_i$ if $(i,j) \in g$.

Given $a \in A$ and $g \in G$, the total payoff a player $i \in N$ gets, $\Pi_i(a,g)$, is the sum of the net payoffs she gets from all her formed links, *i.e.*, the sum of gross payoffs net of the costs she meets. Let $N_i(g) := \{j \in N \backslash \{i\} : (i,j) \in g\}$, then

$$\Pi_i(a,g) := \sum_{j \in N_i(g)} \hat{\pi}(a_i, a_j) = \sum_{j \in N_i(g)} \left(\pi(a_i, a_j) - c(a_i, a_j)\right). \tag{1}$$

2.1 Link Formation and Cost Sharing

Let $a \in A$ be the action profile chosen in the first stage. Let $i, j \in N$, $j \neq i$, be such that $\pi(a_i, a_j) + \pi(a_j, a_i) > c$. If i and j do not form the link (i,j), then they obtain the disagreement value, *i.e.*, $(0,0)$. Differently, if i and j form the link, they obtain $(\hat{\pi}(a_i, a_j), \hat{\pi}(a_j, a_i)) = \mathcal{NS}(B(a_i, a_j), (0,0))$. Since the Nash solution results in a strictly positive sharing of the surplus for each player, in equilibrium the link (i,j) forms. Hence, given a, there is a unique network under which (a,g) can be an equilibrium. We call this network $g(a)$:

$$g(a) = \{(i,j) \in N \times N : i \neq j \text{ and } \pi(a_i, a_j) + \pi(a_j, a_i) > c\} \in G. \tag{2}$$

In the next proposition we identify the sharing of the surplus corresponding to the Nash solution.

Proposition 1 *For all $i, j \in N$ and all $a_i, a_j \in \{\alpha, \beta\}$, if $\pi(a_i, a_j) + \pi(a_j, a_i) > c$, then $\hat{\pi}(a_i, a_j) = \max\left\{\pi(a_i, a_j) - c, \min\left\{\frac{\pi(a_i,a_j)+\pi(a_j,a_i)-c}{2}, \pi(a_i, a_j)\right\}\right\}$.*

Proof. See the appendix. □

Corollary 1 *For all $i, j \in N$ and all $a_i, a_j \in \{\alpha, \beta\}$, if the link (i,j) is formed, then $c(a_i, a_j) = \min\left\{c, \max\left\{\frac{c+\pi(a_i,a_j)-\pi(a_j,a_i)}{2}, 0\right\}\right\}$. Moreover $c(a_i, a_j) + c(a_j, a_i) = c$.*

Proof. The first part follows from substituting $\hat{\pi}(a_i, a_j)$ from Proposition 1 in $c(a_i, a_j) = \pi(a_i, a_j) - \hat{\pi}(a_i, a_j)$. The fact that $c(a_i, a_j) + c(a_j, a_i) = c$ is immediate. □

Note that, if both players have the same gross payoff, then they share the cost evenly. We observe that $c(a_i, a_j)$ is (weakly) increasing in $\pi(a_i, a_j) - \pi(a_j, a_i)$. Thus, given a link (i,j), the player with the highest gross payoff meets a higher share of the cost and, the higher the difference of gross payoffs, the higher the difference in

[9] Recall that we have explicitly assumed that, if $\pi(a_i, a_j) + \pi(a_j, a_i) \leq c$, then $j \notin S^i(a)$.

the cost shares. In particular, when $\pi(a_i, a_j) - \pi(a_j, a_i) > c$, player i meets the entire cost.

Note that, for all $i, j \in N$, $j \neq i$ and all $a \in A$,

$$\hat{\pi}(a_i, a_j) = \begin{cases} 0 & \text{if } \pi(a_i, a_j) + \pi(a_j, a_i) \leq c \\ \max\left\{\pi(a_i, a_j) - c, \min\{\frac{\pi(a_i,a_j)+\pi(a_j,a_i)-c}{2}, \pi(a_i, a_j)\}\right\} & \text{otherwise.} \end{cases} \tag{3}$$

Hence, by equation (2), for all $i \in N$, $N_i(g(a)) = \{j \in N \backslash \{i\} : \pi(a_i, a_j) + \pi(a_j, a_i) > c\}$ and $\Pi_i(a, g(a)) = \sum_{j \in N_i(g(a))} \hat{\pi}(a_i, a_j) = \sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i, a_j)$. Table 2 represents $\hat{\pi}(a_i, a_j)$ separating the cases $c \leq f - e$ and $c > f - e$.

	α	β
α	$\frac{2d-c}{2}$	e
β	$f - c$	$\max\{\frac{2b-c}{2}, 0\}$

$c \leq f - e$

	α	β
α	$\frac{2d-c}{2}$	$\max\{\frac{e+f-c}{2}, 0\}$
β	$\max\{\frac{e+f-c}{2}, 0\}$	$\max\{\frac{2b-c}{2}, 0\}$

$c > f - e$

Table 2: the bilateral net payoff matrix.

Lemma 1 *$\hat{\pi}(\alpha, \alpha) - \hat{\pi}(\beta, \alpha) > 0$ and $\hat{\pi}(\alpha, \alpha) + \hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta) + \hat{\pi}(\beta, \alpha)$.*

Proof. Since $d > f$ and $d > b > e$, $\hat{\pi}(\alpha, \alpha) - \hat{\pi}(\beta, \alpha) > 0$. Note that $\hat{\pi}(\alpha, \alpha) = d - c/2 > 0$, $\hat{\pi}(\beta, \beta) = \max\{b - c/2, 0\}$ and $\hat{\pi}(\alpha, \beta) + \hat{\pi}(\beta, \alpha) = \max\{e + f - c, 0\}$. Hence, if $c \geq e + f$, $\hat{\pi}(\alpha, \alpha) + \hat{\pi}(\beta, \beta) \geq d - c/2$ and $\hat{\pi}(\alpha, \beta) + \hat{\pi}(\beta, \alpha) = 0$. If $c < e + f$, then $\hat{\pi}(\alpha, \beta) + \hat{\pi}(\beta, \alpha) = e + f - c$. Now there are two possibilities. First, if $c \leq 2b$, then $\hat{\pi}(\alpha, \alpha) + \hat{\pi}(\beta, \beta) = d + b - c > e + f - c$. Second, if $c > 2b$, $\hat{\pi}(\alpha, \alpha) + \hat{\pi}(\beta, \beta) = d - c/2 > e + f - c$. □

Note that, by Lemma 1, (α, α) is a strict Nash equilibrium of the (symmetric) bilateral game represented in Table 2.

2.2 Equilibrium analysis

Let $A^* := \{a \in A : (a, g(a))$ is an equilibrium$\}$ and let $a^\alpha, a^\beta \in A$ be defined, for all $i \in N$, by $a_i^\alpha := \alpha$ and $a_i^\beta := \beta$.

Lemma 2 *If $a^* \in A^*$, then, for all $i, j \in N$, $a_i^* = a_j^*$.*

Proof. See the appendix. □

Corollary 2 *For all $a^* \in A^*$, if $(i, j) \in g(a^*)$, then $c(a_i, a_j) = c(a_j, a_i) = \frac{c}{2}$.*

Proof. The result follows from combining Lemma 2 and Corollary 1. □

Proposition 2 *(I) If $2b < e + f$, $A^* = \begin{cases} \{a^\alpha\} & \text{if } c \in (2(b-e), f+e) \\ \{a^\alpha, a^\beta\} & \text{otherwise.} \end{cases}$*
(II) If $2b \geq e + f$, $A^* = \{a^\alpha, a^\beta\}$.

Proof. See the appendix. □

By equation (2) and Proposition 2 we can characterize the network associated to every equilibrium. Let $g^E \in G$ denote the empty network, *i.e.*, $g^E = \emptyset$, and let $g^C \in G$ denote the complete network, *i.e.*, for all $i, j \in N, i \neq j$, $(i, j) \in g^C$.

Corollary 3 $g(a^\alpha) = g^C$ *and, if* $a^\beta \in A^*$, $g(a^\beta) = \begin{cases} g^E & \text{if } c \geq \max\{2b, f+e\} \\ g^C & \text{otherwise.} \end{cases}$

Proof. By equation (2), since $2d > c$, $g(a^\alpha) = g^C$. Now, note that $g(a^\beta) = g^C$ if and only if $2b > c$. By Proposition 2, if $2b \geq e + f$, then $a^\beta \in A^*$ and, if $2b < e + f$, then $a^\beta \in A^*$ if and only if $c \geq f + e$. □

We have shown that there exists a wide range of parameter values for which both a^α and a^β are equilibria and, in some cases, both the complete and the empty network are equilibrium networks. To refine the set of equilibria, we now introduce and analyze a dynamic version of the model.

3 The Dynamic Model

Time is considered to be discrete, $t = 0, 1, 2, ...$ At each period t, the state of the system consists of an action profile $a(t) \in A$ and a network $g(t) \in G$. Let $(a(0), g(0))$ denote the initial state. As in the static case, for all $i, j \in N$, $j \neq i$, we assume that, if $\pi(a_i(0), a_j(0)) + \pi(a_j(0), a_i(0)) \leq c$, then $(i, j) \notin g(0)$ and that, if $(i, j) \in g(0)$, then $\hat{\pi}(a_i(0), a_j(0)) = \mathcal{NS}(B(a_i(0), a_j(0)), (0, 0))$. At each period $t \geq 1$ there are two stages:

First stage. With probability $p \in (0, 1)$, each player $i \in N$ independently receives an opportunity to change her action. Thus, the probability that only the players in $\{i_1, i_2, ..., i_z\} \subseteq N$ receive a revision opportunity at period t is $p^z \cdot (1 - p)^{n-z}$. If $i \in N$ receives a revision opportunity, she can choose a new action $a_i(t) \in \{\alpha, \beta\}$ and, otherwise, $a_i(t) = a_i(t - 1)$. We assume that all players that receive a revision opportunity at period t make their choices simultaneously.

Second stage. Once the action profile $a(t)$ has been chosen, it is observed by all the population. Then, for each pair $i, j \in N$, $j \neq i$, if at least one of the two players received a revision opportunity, the link (i, j) is revised; the link formation being now as in the static case (Subsection 2.1). If neither player i nor j received a revision opportunity in the first stage, then the link is unchanged, *i.e.*, $(i, j) \in g(t) \iff (i, j) \in g(t - 1)$ and $\hat{\pi}(a_i(t), a_j(t)) = \hat{\pi}(a_i(t - 1), a_j(t - 1))$.

For all $i \in N$, the (total) payoff at period t is computed as in the static case, *i.e.*, $\Pi_i(a(t), g(t)) = \sum_{j \in N_i(g(t))} \hat{\pi}(a_i(t), a_j(t))$.

Remark 1 *If $i \in N$ receives a revision opportunity at period $\bar{t}$, then, for all $t \geq \bar{t}$ and all $j \in N \backslash \{i\}$, $(i,j) \in g(t) \Leftrightarrow (i,j) \in g(a(t))$, where $g(a(t))$ is defined by (2). Hence, for all $t \geq \bar{t}$, $N_i(g(t)) = N_i(g(a(t)))$ and $\Pi_i(a(t), g(t)) = \Pi_i(a(t), g(a(t))) = \sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i(t), a_j(t))$, where $\hat{\pi}(a_i, a_j)$ is defined by equation (3).*

We assume that each player that revises her action chooses a myopic best response to the actions that where chosen by the remaining players in the previous period, *i.e.*, when choosing $a_i(t)$, player i believes that the bargaining outcome of the second stage will be determined on the basis of $(a_i(t), a_{-i}(t-1))$. Hence, in view of Remark 1, for all $i \in N$ and all $t \geq 1$, with probability $1-p$, $a_i(t) = a_i(t-1)$ and, with probability p, $a_i(t) \in \arg\max_{a_i \in \{\alpha,\beta\}} \sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i, a_j(t-1))$.[10] If the later arg max is $\{\alpha, \beta\}$, then there is $p_\alpha \in (0,1)$ such that, with probability p_α, $a_i(t) = \alpha$ and, with probability $(1-p_\alpha)$, $a_i(t) = \beta$.[11] Note that, once all the players that received a revision opportunity at period t have (myopically) chosen their actions, the realized action profile $a(t)$ is observed by all the population. Then, in the second stage, the network and cost sharing are updated according to $a(t)$. Hence, the payoff at period t of a player $i \in N$ that received a revision opportunity is $\Pi_i(a(t), g(t)) = \sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i(t), a_j(t))$, which may differ from $\sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i(t), a_j(t-1))$.

Remark 2 *We intend to characterize the limit states of the dynamics. In view of Remark 1, the dynamics is governed by the players' choices of actions. To simplify the exposition, we assume that $g(0) = g(a(0))$. This can be done without loss of generality, since, if $g(t) \neq g(a(t))$, $\Pr(g(t+1) = g(a(t+1))) > 0$ and, if $g(t) = g(a(t))$, for all $\tau > t$, $g(\tau) = g(a(\tau))$. Hence, for every limit state (a, g), $g = g(a)$ (equation (2)).*

Given the arguments presented in Remarks 1 and 2, we can reduce the state space to A, *i.e.*, the state at period t is $a(t) \in A$. For every $a(t) \in A$, by equation (2) and Corollary 1, we can uniquely identify a network, $g(a(t))$, and a profile of cost shares as in the static setting.

3.1 Limit sets and limit states

Let $\Delta(A)$ be the set of probability distributions over A. If we assume that $a(0)$ is chosen trough a certain $\mu_0 \in \Delta(A)$, the dynamics described above determines a Markov chain on A. Let Q represent the $2^n \times 2^n$ transition matrix, where $(Q)_{a,\hat{a}} := \Pr(a(t) = \hat{a} \mid_{a(t-1)=a})$ for all $a, \hat{a} \in A$. Then, the probability that the Markov process (A, Q) leads to each state at period $t \geq 1$ is $\mu_t = \mu_0 \cdot Q^t$. Our first objective is to characterize the limit sets of the process (A, Q). Formally, a set $L \subset A$ is a limit set of (A, Q) if, for all $\hat{a} \in L$, $\sum_{a \in L} (Q)_{\hat{a},a} = 1$ and, for all $\hat{a}, \tilde{a} \in L$, there exists $t > 0$

[10] The results of the paper would hold if we considered that players use a (myopic) better response, *i.e.*, $a_i(t) \in \{\bar{a}_i \in \{\alpha, \beta\} : \sum_{j \in N \backslash \{i\}} \hat{\pi}(\bar{a}_i, a_j(t-1)) \geq \sum_{j \in N \backslash \{i\}} \hat{\pi}(a_i(t-1), a_j(t-1))\}$, as long as all the better responses are chosen with positive probability.

[11] The results of the paper hold if we allow p_α to differ across players and/or across states.

such that $(Q^t)_{\hat{a},\tilde{a}} > 0$. Let $\mathcal{L}$ be the set formed by all the limit sets of (A, Q). A state $a \in A$ is a limit state if it belongs to some limit set. Let A^l be the set of limit states, *i.e.*, $A^l = \cup_{L \in \mathcal{L}} L$.

We now introduce some constants that will be useful throughout the analysis. Let $\phi := \frac{\hat{\pi}(\beta,\beta)-\hat{\pi}(\alpha,\beta)}{\hat{\pi}(\beta,\beta)-\hat{\pi}(\alpha,\beta)+\hat{\pi}(\alpha,\alpha)-\hat{\pi}(\beta,\alpha)}$ and let $\kappa := \phi \cdot (n-1)$. By Lemma 1, $\phi < 1$ and $\kappa < n-1$. In the next lemma we show that, for each $x \in \{\alpha, \beta\}$, $\{a^x\}$ constitutes a limit set of the dynamics if and only if (x, x) is a strict Nash equilibrium of the bilateral game represented in Table 2. Moreover, another consequence is that each limit set is a singleton, *i.e.*, a limit state.

Lemma 3 $\mathcal{L} = \begin{cases} \{\{a^\alpha\}\} & \textit{if} \quad \hat{\pi}(\beta,\beta) \leq \hat{\pi}(\alpha,\beta) \\ \{\{a^\alpha\}, \{a^\beta\}\} & \textit{if} \quad \hat{\pi}(\beta,\beta) > \hat{\pi}(\alpha,\beta). \end{cases}$

Proof. We analyze, for each $a \in A$, the one-period transitions from a. Given $i \in N$, $a_i \rightsquigarrow \alpha$ denotes that player i chooses α in the next period (similarly for $a_i \rightsquigarrow \beta$). Let $K_a = \{i \in N : a_i = \alpha\}$ and $k_a = |K_a|$.

Let $i \in K_a$. Then, $\sum_{l \in N \setminus \{i\}} \hat{\pi}(\alpha, a_l) = (k_a - 1) \cdot \hat{\pi}(\alpha,\alpha) + (n - k_a) \cdot \hat{\pi}(\alpha,\beta)$ and $\sum_{l \in N \setminus \{i\}} \hat{\pi}(\beta, a_l) = (k_a - 1) \cdot \hat{\pi}(\beta,\alpha) + (n - k_a) \cdot \hat{\pi}(\beta,\beta)$. Hence, by the dynamics, with probability $1-p$, $a_i \rightsquigarrow \alpha$ and, with probability p, player i is given a chance to revise. If so, since, by Lemma 1, $\hat{\pi}(\alpha,\alpha) + \hat{\pi}(\beta,\beta) > \hat{\pi}(\alpha,\beta) + \hat{\pi}(\beta,\alpha)$, $a_i \rightsquigarrow \alpha$ if $k_a > \kappa + 1$ and also, with probability p_α, if $k_a = \kappa + 1$; otherwise $a_i \rightsquigarrow \beta$. Let $j \in N \setminus K_a$. Then, $\sum_{l \in N \setminus \{j\}} \hat{\pi}(\beta, a_l) = k_a \cdot \hat{\pi}(\beta,\alpha) + (n - k_a - 1) \cdot \hat{\pi}(\beta,\beta)$ and $\sum_{l \in N \setminus \{j\}} \hat{\pi}(\alpha, a_l) = k_a \cdot \hat{\pi}(\alpha,\alpha) + (n - k_a - 1) \cdot \hat{\pi}(\alpha,\beta)$. Hence, by the dynamics, with probability $1-p$, $a_j \rightsquigarrow \beta$ and, with probability p, player j is given a chance to revise. If so, since $\hat{\pi}(\alpha,\alpha) + \hat{\pi}(\beta,\beta) > \hat{\pi}(\alpha,\beta) + \hat{\pi}(\beta,\alpha)$, $a_j \rightsquigarrow \alpha$ if $k_a > \kappa$ and also, with probability p_α, if $k_a = \kappa$; otherwise $a_j \rightsquigarrow \beta$.

We can already determine the transition matrix Q. Nevertheless, we only need to identify some aspects of (A, Q). Given $a \in A$, the behavior of a player in K_a that revises her action depends on whether $k_a > \kappa + 1$, $k_a = \kappa + 1$ or $k_a < \kappa + 1$; and the behavior of a player in $N \setminus K_a$ that revises her action depends on whether $k_a > \kappa$, $k_a = \kappa$ or $k_a < \kappa$. We distinguish the following possibilities: (i) $\kappa \in (0, n-1) \cap \mathbb{Z}$, (ii) $\kappa \in (0, n-1) \setminus \mathbb{Z}$, (iii) $\kappa = 0$, (iv) $\kappa < 0$. The dynamics is represented in Figure 1, where the numbers inside the boxes indicate the possible values that k_a can take. A path of consecutive arrows starting at box $m \in \{0, 1, ..., n\}$ and ending at box $m' \in \{0, 1, ..., n\}$ means that, for all $a \in A$ such that $k_a = m$, there exists $a' \in A$ such that $k_{a'} = m'$ and $(Q)_{a,a'} > 0$,[12] *i.e.*, there is positive probability of (direct) transition from a state with m players playing α to a state with m' players playing α. The absence of a path of consecutive arrows from m to m' means that, for all $a, a' \in A$ such that $k_a = m$ and $k_{a'} = m'$, $(Q)_{a,a'} = 0$.

[12] In Figure 1, two arrows are consecutive if the first arrow ends in the same box that the second arrow begins. Note that, if no player receives a revision opportunity at t, $a(t) = a(t-1)$. Hence, for all $a \in A$, $(Q)_{a,a} \geq (1-p)^n > 0$.

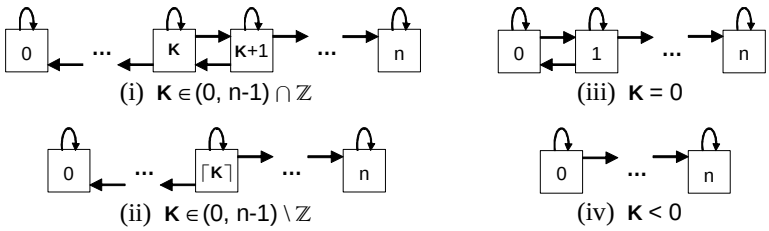

Figure 1. The dynamics for different values of κ.
$\lceil \kappa \rceil$ denotes the least integer not less than κ.

We formally prove case (i) and the rest of the cases follow by analogous reasoning. Let $\kappa \in (0, n-1) \cap \mathbb{Z}$. We distinguish four cases: (i.1) $k_a < \kappa$, (i.2) $k_a = \kappa$, (i.3) $k_a = \kappa + 1$, and (i.4) $k_a > \kappa + 1$.

(i.1) $k_a < \kappa$. Then, $k_a < \kappa + 1$ and, for every $i \in N$ that receives a revision opportunity, $a_i \rightsquigarrow \beta$. Hence, for all $m \in \{0, 1, ..., k_a\}$, there exist $\tilde{a} \in A$ such that $k_{\tilde{a}} = m$ and $(Q)_{a,\tilde{a}} > 0$. Moreover, for all $\hat{a} \in A$ such that $k_{\hat{a}} > k_a$, $(Q)_{a,\hat{a}} = 0$.

(i.2) $k_a = \kappa$. Now, for each $i \in K_a$ that is given the chance to revise, $a_i \rightsquigarrow \beta$ and, for each $j \in N \backslash K_a$ that is given the chance to revise, with probability p_α, $a_j \rightsquigarrow \alpha$ and, with probability $1 - p_\alpha$, $a_j \rightsquigarrow \beta$. Hence, for all $m \in \{0, 1, ..., n\}$, there exist $\tilde{a} \in A$ such that $k_{\tilde{a}} = m$ and $(Q)_{a,\tilde{a}} > 0$.

Cases (i.3) and (i.4) are analogous to (i.2) and (i.1), respectively, but exchanging the roles of α and β.

An examination of Figure 1 shows that, in cases (iii)-(iv), *i.e.*, if $\kappa \leq 0$, $\mathcal{L} = \{\{a^\alpha\}\}$ and, in cases (i)-(ii), *i.e.*, if $\kappa \in (0, n-1)$, $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$. By Lemma 1, if $\hat{\pi}(\beta, \beta) \leq \hat{\pi}(\alpha, \beta)$, $\kappa \leq 0$ and, if $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$, $\kappa \in (0, n-1)$. □

Proposition 3 $\mathcal{L} = \begin{cases} \{\{a^\alpha\}, \{a^\beta\}\} & \textit{if } 2b > e + f \textit{ and } c < 2b \\ \{\{a^\alpha\}, \{a^\beta\}\} & \textit{if } 2b \leq e + f \textit{ and } c < 2(b - e) \\ \{\{a^\alpha\}\} & \textit{otherwise.} \end{cases}$

Proof. See the appendix. □

By comparing Proposition 2 and Proposition 3, we can see that the set of limit states is a subset (maybe strict) of A^*, *i.e.*, the dynamics allows to refine the set of equilibria. The next corollary characterizes the networks associated to the limit states and the costs met by the players.

Corollary 4 *For all $a \in A^l$, $g(a) = g^C$, and for all $(i, j) \in g(a)$, $c(a_i, a_j) = c/2$.*

Proof. Since $A^l \subseteq A^*$, by Corollary 3, $g(a) \in \{g^C, g^E\}$ and, moreover, g^E only appears (as $g(a^\beta)$) if $\max\{2b, f + e\} < c$. However, by Proposition 3, in such a case, $A^l = \{a^\alpha\}$ and $g(a^\alpha) = g^C$. Since $A^l \subseteq \{a^\alpha, a^\beta\}$, for all $a \in A^l$ and all $(i, j) \in g(a)$, $\pi(a_i, a_j) = \pi(a_j, a_i)$. Hence, by Corollary 1, $c(a_i, a_j) = c(a_j, a_i) = c/2$. □

3.2 Perturbed dynamics: stochastically stable states

Now, in order to further refine the set of equilibria, we introduce a perturbation in the dynamics and study the set of stochastically stable states. We follow the approach in Ellison (2000). More specifically, we allow for mutations, interpreted as the possibility of players making errors when implementing their choices on actions, *i.e.*, at all period $t \geq 1$, each player that receives a revision opportunity will, with a small probability $\varepsilon > 0$ revise her action randomly; namely, with probability $q_\alpha \in (0,1)$ chooses α and with probability $1 - q_\alpha$ chooses β.[13] The mutations are produced independently across players.

Remark 3 *If, additionally to the mutations in the choices of actions, we considered mutations in the formation of the links,*[14] *the result of the paper would be unaltered, i.e., there is no combination of mutations of the links that allow for transitions among different limit sets. This is due to the fact that, throughout the dynamics, after a player receives an opportunity to revise her action all her links are also revised on the basis of her new action. Hence, mutations in the formation of the links would not alter the posterior players' choices of actions.*

Specifically, the perturbed dynamics is as follows: for all $i \in N$ and all $t \geq 1$, with probability $1 - p$, $a_i(t) = a_i(t-1)$; with probability $p \cdot \varepsilon \cdot q_\alpha$, $a_i(t) = \alpha$; with probability $p \cdot \varepsilon \cdot (1 - q_\alpha)$, $a_i(t) = \beta$ and, with probability $p \cdot (1 - \varepsilon)$, $a_i(t) \in \arg\max_{a_i \in \{\alpha,\beta\}} \sum_{j \in N \setminus \{i\}} \hat{\pi}(a_i, a_j(t-1))$.[15] If the later arg max is $\{\alpha, \beta\}$, then, with probability p_α, $a_i(t) = \alpha$ and, with probability $(1 - p_\alpha)$, $a_i(t) = \beta$.

Recall that all randomizations are independent across players. For all $\varepsilon > 0$, the process defines a Markov chain on A, with transition matrix Q_ε. Given a Markov process (A, Q_ε), $\mu \in \Delta A$ is an invariant probability distribution if $\mu \cdot Q_\varepsilon = \mu$. Since for all $a, a' \in A$, $(Q_\varepsilon)_{a,a'} > 0$, (A, Q_ε) has a unique invariant probability distribution, namely μ_ε.[16] From Ellison (2000), it is also known that $\tilde{\mu} := \lim_{\varepsilon \to 0} \mu_\varepsilon$ is well defined. A state $a \in A$ is stochastically stable when $\tilde{\mu}(a) > 0$. Let $A^s := \{a \in A : \tilde{\mu}(a) > 0\}$. Young (1993) shows that, if a state is stochastically stable, it is a limit state of the unperturbed dynamics, *i.e.*, $A^s \subseteq A^l$.

We define a cost function $C : A \times A \to \mathbb{N} \cup \{0\}$ such that, for all pairs $a, \hat{a} \in A$, $0 < \lim_{\varepsilon \to 0} (Q_\varepsilon)_{a,\hat{a}} / \varepsilon^{C(a,\hat{a})} < \infty$, *i.e.*, $(Q_\varepsilon)_{a,\hat{a}}$ and $\varepsilon^{C(a,\hat{a})}$ are infinitessimals in ε of the same order. For all $a, \hat{a} \in A$, $C(a, \hat{a})$ represents the minimum number of mutations needed for a transition from a to $\hat{a}$ to be feasible. Note that if a transition from a to $\hat{a}$ has positive probability under the unperturbed dynamics, *i.e.*, if $(Q_0)_{a,\hat{a}} \equiv (Q)_{a,\hat{a}} > 0$, then $C(a, \hat{a}) = 0$.

[13] Again, the results of the paper hold if we allow q_α to differ across players and/or across states.

[14] For instance, that at any period $t \geq 1$, once the network $g(t)$ has formed, each link $(i,j) \in g(t)$ is severed with a small probability $\delta > 0$.

[15] Note that, if we considered $p = 1$, we would obtain the dynamics of Kandori *et al.* (1993). In such a case the results of the paper essentially hold.

[16] See, for instance, Freidlin and Wentzell (1984).

Below we introduce some concepts from Ellison (2000). A path from $Z \subseteq A$ to a set $Y \subseteq A \backslash Z$ is a finite sequence of distinct states $\left(a^1, ..., a^V\right)$ with $a^1 \in Z$, $a^v \notin Y$ for $2 \leq v \leq V-1$, and $a^V \in Y$. Let $S(Z, Y)$ to be the set of all paths from Z to Y. We can extend the definition of cost to paths of states by setting $C(a^1, ..., a^V) = \sum_{u=1}^{V-1} C(a^u, a^{u+1})$. Let $\hat{L}$ be a union of limit sets of the unperturbed Markov process (A, Q). The basin of attraction of $\hat{L}$, denoted $D(\hat{L})$, is the set of initial states from which (A, Q) converges to $\hat{L}$ with probability one. Now we define the radius and coradius of the basin of attraction of a union of limit sets of the unperturbed process $\hat{L}$, which are denoted by $R(\hat{L})$ and $CR(\hat{L})$, respectively.

$$\begin{aligned} R(\hat{L}) &= \min_{(a^1,...,a^V) \in S(\hat{L}, A \backslash D(\hat{L}))} C(a^1, a^2, ..., a^V). \\ CR(\hat{L}) &= \max_{a \in A \backslash \hat{L}} \min_{(a^1,...,a^V) \in S(\{a\}, \hat{L})} C(a^1, a^2, ..., a^V). \end{aligned}$$

Intuitively, the radius of a set $\hat{L}$ indicates the minimum cost (minimum number of mutations) required to move out of it, and the coradius is the maximum cost required to move into this set from some limit state not included in it.

Theorem 1 ***(Ellison (2000)).*** *If for some set $\hat{L}$, which is a union of limit sets, $R(\hat{L}) > CR(\hat{L})$, then $A^s \subseteq \hat{L}$.*

In order to identify the set of stochastically stable states (Theorem 2), we first state and prove Lemmas 4 and 5.

Lemma 4 *Assume $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$.*
(I) If n is odd, then $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa > \frac{n-1}{2}$ and $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa < \frac{n-1}{2}$.
(II) If n is even, then $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa \geq \frac{n}{2}$ and $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa \leq \frac{n}{2} - 1$.

Proof. By Lemma 3, $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$ implies $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$. Also, since $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$ and $\hat{\pi}(\alpha, \alpha) > \hat{\pi}(\beta, \alpha)$, $\kappa \in (0, n-1)$. We distinguish two cases:
i) $\kappa \in (0, n-1) \cap \mathbb{Z}$ (Figure 1 (i)). The basins of attraction of the two limit sets $\{a^\alpha\}$ and $\{a^\beta\}$ are $D(\{a^\beta\}) = \{a \in A : k_a \leq \kappa - 1\}$ and $D(\{a^\alpha\}) = \{a \in A : k_a \geq \kappa + 2\}$. In Figure 2 (i) we represent the perturbed dynamics when $\kappa \in (0, n-1) \cap \mathbb{Z}$. For all $\varepsilon > 0$ and all $a, a' \in A$, $(Q_\varepsilon)_{a,a'} > 0$, which is represented by the fact that, for all $a, a' \in A$, there exists a path of consecutive arrows from k_a to $k_{a'}$. The regular arrows represent transitions that do not require mutations and, therefore imply cost 0. On the other hand, a dashed arrow from one box to another indicates that a mutation is needed from the corresponding transition to take place. The minimum number of mutations needed for a transition from a state with m players choosing α to a state with m' choosing α, *i.e.*, $\min_{(a^1,...,a^V) \in S(\{\hat{a} \in A : k_{\hat{a}} = m\}, \{\tilde{a} \in A : k_{\tilde{a}} = m'\})} C(a^1, ..., a^V)$, is represented by the minimum number of dashed arrows among all paths of consecutive arrows starting at box m and ending at box m'.

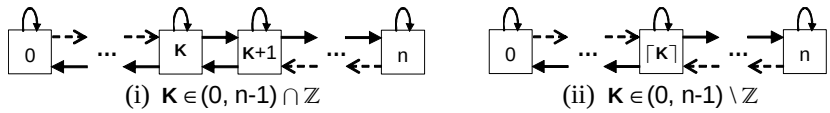

Figure 2. The perturbed dynamics when $\kappa \in (0, n-1)$.

An examination of Figure 2 (i) indicates that the minimum number of mutations needed for a transition from a^β ($k_{a^\beta} = 0$) to a state in $A \backslash D(\{a^\beta\})$ is $C(a^\beta, \hat{a})$, where $k_{\hat{a}} = \kappa$. The minimum number of mutations needed for a transition from a^α ($k_{a^\alpha} = n$) to a state in $A \backslash D(\{a^\alpha\})$ is $C(a^\alpha, \tilde{a})$, where $k_{\tilde{a}} = \kappa + 1$. Thus, $R(\{a^\beta\}) = CR(\{a^\alpha\}) = \kappa$ and $CR(\{a^\beta\}) = R(\{a^\alpha\}) = n - \kappa - 1$. Hence, if $\kappa \in (0, n-1) \cap \mathbb{Z}$, then

$$R(\{a^\beta\}) > CR(\{a^\beta\}) \text{ if and only if } \kappa > (n-1)/2, \text{ and} \tag{4}$$
$$R(\{a^\alpha\}) > CR(\{a^\alpha\}) \text{ if and only if } \kappa < (n-1)/2. \tag{5}$$

ii) $\kappa \in (0, n-1) \backslash \mathbb{Z}$ (Figure 1 (ii)). The basins of attraction of two limit sets are $D(\{a^\beta\}) = \{a \in A : k_a \leq \lceil \kappa \rceil - 1\}$ and $D(\{a^\alpha\}) = \{a \in A : k_a \geq \lceil \kappa \rceil + 1\}$, where $\lceil \kappa \rceil$ denotes the least integer not less than κ. In Figure 2 (ii), analogously to the previous case, we represent the perturbed dynamics when $\kappa \in (0, n-1) \backslash \mathbb{Z}$. An examination of Figure 2 (ii) indicates that the minimum number of mutations needed for a transition from a^β ($k_{a^\beta} = 0$) to a state in $A \backslash D(\{a^\beta\})$ is $C(a^\beta, \hat{a})$, where $k_{\hat{a}} = \lceil \kappa \rceil$. The minimum number of mutations needed for a transition from a^α ($k_{a^\alpha} = n$) to a state in $A \backslash D(\{a^\alpha\})$ is $C(a^\alpha, \hat{a})$, where $k_{\hat{a}} = \lceil \kappa \rceil$. Thus, $R(\{a^\beta\}) = CR(\{a^\alpha\}) = \lceil \kappa \rceil$ and $CR(\{a^\beta\}) = R(\{a^\alpha\}) = n - \lceil \kappa \rceil$. Hence, if $\kappa \in (0, n-1) \backslash \mathbb{Z}$, then

$$R(\{a^\beta\}) > CR(\{a^\beta\}) \text{ if and only if } \lceil \kappa \rceil > n/2, \text{ and} \tag{6}$$
$$R(\{a^\alpha\}) > CR(\{a^\alpha\}) \text{ if and only if } \lceil \kappa \rceil < n/2. \tag{7}$$

We now group the results obtained in i) and ii) and characterize the conditions for $R(\{a^\beta\}) > CR(\{a^\beta\})$ and the conditions for $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ when $\kappa \in (0, n-1)$. Since $\lceil \kappa \rceil \in \mathbb{Z}$, we distinguish two cases:

(I) $\kappa \in (0, n-1)$ and n odd. Then $(n-1)/2 \in \mathbb{Z}$ and $n/2 \notin \mathbb{Z}$. Hence, $\lceil \kappa \rceil > n/2$ if and only if $\kappa > (n-1)/2$, and $\lceil \kappa \rceil < n/2$ if and only if $\kappa \leq (n-1)/2$. Therefore, we get the following two results. First, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa \in ((n-1)/2, n-1)$, since, by (4), for all $\kappa \in (0, n-1) \cap \mathbb{Z}$, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa > (n-1)/2$; and, by (6), for all $\kappa \in (0, n-1) \backslash \mathbb{Z}$, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa > (n-1)/2$. Second, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa \in (0, (n-1)/2)$, since, by (5), for all $\kappa \in (0, n-1) \cap \mathbb{Z}$, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa < (n-1)/2$; and, by (7), for all $\kappa \in (0, n-1) \backslash \mathbb{Z}$, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa < (n-1)/2$.

(II) $\kappa \in (0, n-1)$ and n even. Then $n/2 \in \mathbb{Z}$. Hence, $\lceil \kappa \rceil > n/2$ if and only if $\kappa > n/2$, and $\lceil \kappa \rceil < n/2$ if and only if $\kappa \leq n/2 - 1$. Therefore, we get the following two results. First, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa \in [n/2, n-1)$, since, by (4), for all $\kappa \in (0, n-1) \cap \mathbb{Z}$, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa \geq n/2$;

and, by (6), for all $\kappa \in (0, n-1)\backslash\mathbb{Z}$, $R(\{a^\beta\}) > CR(\{a^\beta\})$ if and only if $\kappa > n/2$. Second, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa \in (0, n/2-1]$, since, by (5), for all $\kappa \in (0, n-1) \cap \mathbb{Z}$, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa \leq n/2-1$; and, by (7), for all $\kappa \in (0, n-1)\backslash\mathbb{Z}$, $R(\{a^\alpha\}) > CR(\{a^\alpha\})$ if and only if $\kappa < n/2-1$. □

Lemma 5 *(I) Assume n is odd. If $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) > \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$, $A^s = \{a^\beta\}$ and, if $\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) > \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta)$, $A^s = \{a^\alpha\}$.*
(II) Assume n is even. If $(\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-2) \geq (\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))n$, $A^s = \{a^\beta\}$ and, if $(\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))(n-2) \geq (\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))n$, $A^s = \{a^\alpha\}$.

Proof. (I) Consider n odd. If $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) > \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$ then $\kappa \in (\frac{n-1}{2}, n-1)$. Hence, by Lemma 4 and Theorem 1, $A^s = \{a^\beta\}$. If $\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) > \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta)$ then $\kappa < \frac{n-1}{2}$. There are two possibilities. If $\kappa \in (0, \frac{n-1}{2})$, then, by Lemma 4 and Theorem 1, $A^s = \{a^\alpha\}$. If $\kappa \leq 0$, then, by Lemma 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$.

(II) Consider n even. First, $(\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-2) \geq (\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))n$ can be rewritten as $2(\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-1) \geq (\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) + \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))n$. The latter inequality becomes $\kappa \geq n/2$, since, by Lemma 1, $\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) + \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) > 0$. Hence, $\kappa \in [n/2, n-1)$ and, by Lemma 4 and Theorem 1, $A^s = \{a^\beta\}$. Second, $(\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))(n-2) \geq (\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))n$ can be rewritten as $2(\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-1) \leq (\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) + \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-2)$. The latter inequality becomes $\kappa \leq (n-2)/2$, since, by Lemma 1, $\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha) + \hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) > 0$. There are two possibilities. If $\kappa \in (0, \frac{n-2}{2}]$, then, by Lemma 4 and Theorem 1, $A^s = \{a^\alpha\}$. If $\kappa \leq 0$, then, by Lemma 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$. □

Before presenting the main result of the paper, we define a measure of the degree of risk-dominance of action β in the bilateral coordination game (Table 1): the risk-dominance premium, r. It is a particularization of the optimization premium of Battalio *et al.* (2001). The latter is defined as the difference in the payoff of the best response to an opponent's strategy and the worst response. Since action β is risk-dominant, it is the unique best response to an opponent mixing 50/50 in our bilateral game. We define r as the optimization premium in such a case, *i.e.*, $r = (b+f-d-e)/2$.[17] Let $\bar{n} := (2d-e-f)/(d-b)$ and, for all $x > 1$, define the positive strictly decreasing function $\delta(x) = (b-e+d-f)/(2(x-1))$. Note that $\lim_{x\to\infty} \delta(x) = 0$. In the following theorem we identify the set of stochastically stable states.

Theorem 2 *(I) For all n odd, if $\frac{c}{2} < r$, $A^s = \{a^\beta\}$ and, if $\frac{c}{2} > r$, $A^s = \{a^\alpha\}$.*
(II) For all n even higher than $\bar{n}$, if $\frac{c}{2} < r - \delta(n)$, $A^s = \{a^\beta\}$ and, if $\frac{c}{2} > r + \delta(n)$, $A^s = \{a^\alpha\}$.

[17] Intuitively, a player who is uncertain about the action of her opponent, and assigns equal probability of her playing each action, has a higher incentive to play β the higher r is.

Proof. See the appendix. □

The next corollary states that, in all stochastically stable states the complete network forms and all the costs of the links are shared evenly.

Corollary 5 *Let $a \in A^s$. Then, $g(a) = g^C$ and, for all $(i,j) \in g(a)$, $c(a_i, a_j) = c(a_j, a_i) = c/2$.*

Proof. Recall that $A^s \subseteq A^l$. Hence the result follows from Corollary 4. □

If n is odd, Theorem 2 has an intuitive interpretation: when r is exceeds the cost that a player pays in equilibrium, *i.e.*, the degree of risk-dominance of action β is high, there is a wide range of parameters for which the risk-dominant state a^β is the unique stochastically stable state. We therefore find a close relationship between the degree of risk-dominance of β and the robustness of the limit state a^β to small perturbations in the dynamics. Conversely, if $r < c/2$, the unique stochastically stable state coincides with the efficient equilibrium.

If n is even, we get a similar result, although there exists an interval centered on r, $[r - \delta(n), r + \delta(n)]$, such that, if $c/2$ lies in this interval, then both a^α and a^β may coexist as stochastically stable states. Note that, the larger the population is, the smaller this interval becomes.

4 Side payments and negative payoffs

We now analyze how allowing for side payments or/and negative payoffs affect to the main result of the paper. If we allow for side payments, the new bilateral net payoff matrix is represented in Table 3. To see this, let $W(a_i, a_j)$ be the bargaining set for $i, j \in N$, $i \neq j$, when side payments are allowed. Note that, if a player is no longer restricted to achieve at most her gross payoff of the coordination game from a link, then, for all $a_i, a_j \in \{\alpha, \beta\}$ such that $\pi(a_i, a_j) + \pi(a_j, a_i) > c$, $W(a_i, a_j) = \{(x_i, x_j) \in \mathbb{R}^2 : x_i + x_j \leq \pi(a_i, a_j) + \pi(a_j, a_i) - c\}$. The net payoffs resulting from the Nash solution are $(w(a_i, a_j), w(a_j, a_i)) = \mathcal{NS}(W(a_i, a_j), (0,0))$, *i.e.*, $w(a_i, a_j) = w(a_i, a_j) = (\pi(a_i, a_j) + \pi(a_j, a_i) - c)/2$ (see proof of Proposition 1).

Under negative payoffs ($e < 0$), the bilateral net payoff matrix is represented in Table 4. To see this, let $a_i = \alpha$ and $a_j = \beta$. Then, $B(a_i, a_j) = \{(x_i, x_j) \in \mathbb{R}^2 : x_i + x_j \leq e + f - c,\ x_i \leq e,\ x_j \leq f\}$. Since the disagreement value is $(0,0)$ and $0 > \hat{\pi}_{i,j}$ for all $\hat{\pi}_{i,j}$ such that $(\hat{\pi}_{i,j}, \hat{\pi}_{j,i}) \in B(a_i, a_j)$, player i does not find profitable to form the link. Hence, no link forms between players choosing different actions.

	α	β
α	$\frac{2d-c}{2}$	$\max\{\frac{e+f-c}{2}, 0\}$
β	$\max\{\frac{e+f-c}{2}, 0\}$	$\max\{\frac{2b-c}{2}, 0\}$

Table 3

	α	β
α	$\frac{2d-c}{2}$	0
β	0	$\max\{\frac{2b-c}{2}, 0\}$

Table 4

Proposition 4 *If side payments are allowed and/or $e < 0$, then (I) for all n odd, $A^s = \{a^\alpha\}$ and, (II) for all n even higher than $\frac{2d-c}{d-b}$, $A^s = \{a^\alpha\}$.*

We omit the proof since it follows similar arguments to the ones used to prove Theorem 2 (using Lemma 5 and considering Tables 3 and 4 instead of Table 2). Note that, if side payments are allowed and/or there are negative payoffs, a^α is the unique stochastically stable state. This implies that the risk-dominant action is never selected in the long run. Clearly, the network associated to the stochastically stable state is complete and the cost of each link is evenly shared between the involved players.

5 Discussion

We compare our results to Goyal and Vega-Redondo (2005), GVR hereafter, and Jackson and Watts (2002), JW hereafter. These two papers also consider the formation of costly links in a context of social coordination. GVR consider a one-sided model, *i.e.*, they assume that players form links on an individual basis: when a player decides to form a link, she unilaterally meets the entire cost of the link. Therefore, as long as the payoffs of the bilateral coordination game are non-negative, no player has incentives to refuse the formation of a link initiated by other player. Differently, JW study a two-sided model, *i.e.*, they assume that, whenever two players form a link, the cost is shared equally, regardless of the actions they choose. Both papers obtain that, when the cost of a link is low, the risk-dominant equilibrium is selected as the unique stochastically stable state. Instead, when cost is high, the efficient equilibrium is selected in GVR, and both the risk-dominant and the efficient equilibria coexist in JW. In both cases, the complete network forms in the long run, as long as the cost of the link is not so high that players do never find profitable to form a link.

Differently to these models, we determine the cost sharing from bargaining. Our bargaining outcome results in symmetric cost sharing when two player coordinate on the same action, and in asymmetric cost sharing when they are miscoordinated. To this respect, our cost sharing is closer to JW if two players coordinate, and to GVR if they do not. Regarding the network that arises in the long run, we obtain the complete network, as both GVR and JW do. Our results on which equilibrium is selected in the long run (efficient *vs.* risk-dominant) are qualitatively similar to GVR, even though the distribution of the cost in equilibrium in our model coincides with the two-sided model. An explanation of this fact may be found by comparing the revision structures considered in each model. GVR assume that, when a player receives a revision opportunity, she can change both her action and her links (on a unilateral basis), being both decisions taken simultaneously. Differently, JW study a dynamics in which the opportunities to revise actions and links are drawn independently, *i.e.*, players may revise actions and not links or *vice versa*. Our revision structure is closer to GVR, since we assume that, if a player revises her action, then all her links are also revised.

Finally, we would like to point out that we have assumed that the cost of a link in our model is independent of the players that form it. Droste *et al.* (2000) propose a two-sided model with a geographic cost of forming links. They analyze the case in which the players are spatially distributed around a circle and assume that the cost of a link is increasing in the geographic distance between the two players. They obtain that the risk-dominant equilibrium is the unique stochastically stable state. We consider that to introduce bargaining considerations in such a setup and in other contexts of network formation is an interesting issue for future research.

6 Appendix

Proof of Proposition 1

Assume $\pi(a_i, a_j) + \pi(a_j, a_i) > c$. First, let $W(a_i, a_j) = \{(x_i, x_j) \in \mathbb{R}^2 : x_i + x_j \leq \pi(a_i, a_j) + \pi(a_j, a_i) - c\}$. We define the *unrestricted bargaining outcome* as $(w(a_i, a_j), w(a_j, a_i)) = \mathcal{NS}(W(a_i, a_j), (0, 0)) = \arg\max_{(x_i,x_j)\in W(a_i,a_j), x_i\geq 0, x_j\geq 0} x_i \cdot x_j$. From the first order conditions of the maximization problem we get $w(a_i, a_j) = w(a_i, a_j) = (\pi(a_i, a_j)+\pi(a_j, a_i)-c)/2$. Now, we want to calculate $(\hat{\pi}(a_i, a_j), \hat{\pi}(a_j, a_i)) = \mathcal{NS}(B(a_i, a_j), (0, 0))$. There are two possibilities:

First, if $\pi(a_i, a_j) \leq \pi(a_j, a_i)$, the constraint $x_j \leq \pi(a_j, a_i)$ is not binding. In this case, if $\pi(a_j, a_i) \leq c$, the constraint $x_i \leq \pi(a_i, a_j)$ is not binding either. Hence, $\hat{\pi}(a_i, a_j) = w(a_i, a_j)$. If $\pi(a_j, a_i) > c$, there are two cases: if $\pi(a_i, a_j) \geq \pi(a_j, a_i) - c$, the constraint $x_i \leq \pi(a_i, a_j)$ is not binding and, hence, $\hat{\pi}(a_i, a_j) = w(a_i, a_j)$. Differently, if $\pi(a_i, a_j) < \pi(a_j, a_i) - c$, the constraint $x_i \leq \pi(a_i, a_j)$ is binding and the maximization problem results in the corner solution $\hat{\pi}(a_i, a_j) = \pi(a_i, a_j)$.

Second, if $\pi(a_i, a_j) > \pi(a_j, a_i)$, the constraint $x_i \leq \pi(a_i, a_j)$ is not binding. In this case, if $\pi(a_i, a_j) \leq c$, the constraint $x_j \leq \pi(a_j, a_i)$ is not binding either. Hence, $\hat{\pi}(a_i, a_j) = w(a_i, a_j)$. If $\pi(a_i, a_j) > c$, there are two cases: if $\pi(a_j, a_i) \geq \pi(a_i, a_j) - c$, the constraint $x_j \leq \pi(a_j, a_i)$ is not binding and, hence, $\hat{\pi}(a_i, a_j) = w(a_i, a_j)$. Differently, if $\pi(a_j, a_i) < \pi(a_i, a_j) - c$, the constraint $x_j \leq \pi(a_j, a_i)$ is binding and, hence, $\hat{\pi}(a_i, a_j) = \pi(a_i, a_j) - c$. □

Proof of Lemma 2

For all $a \in A$, let $K_a = \{i \in N : a_i = \alpha\}$ and $k_a = |K_a|$. Suppose that there exists $a \in A^*$ such that $k_a \in \{1, ..., n-1\}$.

Case 1: $c \leq f - e$. When $2(b - e) < c \leq f - e$, β is strictly dominated by α, a contradiction with $k_a < n$. Consider $2(b - e) > c$. Since $a \in A^*$, for all $i \in K_a$, $\sum_{j\in N\setminus\{i\}} \hat{\pi}(\alpha, a_j) \geq \sum_{j\in N\setminus\{i\}} \hat{\pi}(\beta, a_j)$, *i.e.*, $(k_a - 1)\frac{2d-c}{2} + (N - k_a)e \geq (N - k_a)\frac{2b-c}{2} + (k_a - 1)(f - c)$. Hence, $k_a \geq \frac{N(2(b-e)-c)}{2(d-f+b-e))} + \frac{2(d-f)+c}{2(d-f+b-e)}$. Moreover, for all $i \in N\setminus K_a$, $\sum_{j\in N\setminus\{i\}} \hat{\pi}(\beta, a_j) \geq \sum_{j\in N\setminus\{i\}} \hat{\pi}(\alpha, a_j)$, *i.e.*, $(N - k_a - 1)\frac{2b-c}{2} + k_a(f - c) \geq (N - k_a - 1)e + k_a\frac{2d-c}{2}$. Hence $k_a \leq \frac{N(2(b-e)-c)}{2(d-f+b-e)} - \frac{2(b-e)-c}{2(d-f+b-e)}$, a contradiction.

Case 2: $c > f - e$. If $\max\{2b, f+e\} \leq c < 2d$, β is weakly dominated by α, a contradiction with $0 < k_a < n$. Consider $f-e < c < \max\{2b, f+e\}$. On the one hand, if $f+e > 2b$, β is strictly dominated by α, a contradiction with $k_a < n$. On the other hand, if $f+e < 2b$, we have two possibilities: $f-e < c < f+e$ and $f+e < c < 2b$. First, if $f-e < c < f+e$, since, for all $i \in K_a$, $\sum_{j \in N \setminus \{i\}} \hat{\pi}(\alpha, a_j) \geq \sum_{j \in N \setminus \{i\}} \hat{\pi}(\beta, a_j)$, then $k_a \geq \frac{N(2b-(e+f))}{2(d-f+b-e)} + \frac{2d-(e+f)}{2(d-f+b-e)}$ and since, for all $i \in N \setminus K_a$, $\sum_{j \in N \setminus \{i\}} \hat{\pi}(\beta, a_j) \geq \sum_{j \in N \setminus \{i\}} \hat{\pi}(\alpha, a_j)$, then $k_a \leq \frac{N(2b-(e+f))}{2(d-f+b-e)} - \frac{2b-(e+f)}{2(d-f+b-e)}$, a contradiction. Second, if $f+e \leq c < 2b$, since, for all $i \in K_a$, $\sum_{j \in N \setminus \{i\}} \hat{\pi}(\alpha, a_j) \geq \sum_{j \in N \setminus \{i\}} \hat{\pi}(\beta, a_j)$, then $k_a \geq \frac{N(2b-c)}{2(d+b-c)} + \frac{2d-c}{2(d+b-c)}$ and since, for all $i \in N \setminus K_a$, $\sum_{j \in N \setminus \{i\}} \hat{\pi}(\beta, a_j) \geq \sum_{j \in N \setminus \{i\}} \hat{\pi}(\alpha, a_j)$, then $k_a \leq \frac{N(2b-c)}{2(d+b-c)} - \frac{2b-c}{2(d+b-c)}$, a contradiction. □

Proof of Proposition 2

By lemma 2, $A^* \subseteq \{a^\alpha, a^\beta\}$. Let Λ denote the set of Nash equilibria of the bilateral game represented in Table 2. Clearly, for all $x \in \{\alpha, \beta\}$, $a^x \in A^*$ if and only if $(x, x) \in \Lambda$. (I) $2b < e+f$. Since $2b < e+f$, $2(b-e) < f-e$. Therefore, if $c \leq 2(b-e)$, $\Lambda = \{(\alpha, \alpha), (\beta, \beta)\}$, hence $A^* = \{a^\alpha, a^\beta\}$. If $2(b-e) < c < f+e$, $\Lambda = \{(\alpha, \alpha)\}$, hence $A^* = \{a^\alpha\}$. Finally, if $c \geq f+e$, $\Lambda = \{(\alpha, \alpha), (\beta, \beta)\}$, hence $A^* = \{a^\alpha, a^\beta\}$. (II) $2b \geq e+f$. Since $2b \geq e+f$, $\Lambda = \{(\alpha, \alpha), (\beta, \beta)\}$, hence $A^* = \{a^\alpha, a^\beta\}$. □

Proof of Proposition 3

First, assume $2b > e+f$.

If $c \leq f-e$, $\hat{\pi}(\beta, \beta) = (2b-c)/2$ and $\hat{\pi}(\alpha, \beta) = e$. Since $2b > e+f$, $2(b-e) > f-e \geq c$. Hence, $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$ and, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$.

If $f-e < c < e+f$, $\hat{\pi}(\beta, \beta) = (2b-c)/2$ and $\hat{\pi}(\alpha, \beta) = (f+e-c)/2$. Since $2b > e+f$, $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$.

If $e+f \leq c < 2b$, $\hat{\pi}(\beta, \beta) = (2b-c)/2 > 0 = \hat{\pi}(\alpha, \beta)$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$.

If $2b \leq c < 2d$, $\hat{\pi}(\beta, \beta) = \hat{\pi}(\alpha, \beta) = 0$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$.

Second, assume $e+f \geq 2b$.

If $c < 2(b-e)$, $\hat{\pi}(\beta, \beta) = (2b-c)/2$ and $\hat{\pi}(\alpha, \beta) = e$, since $e+f \geq 2b$ implies $2(b-e) \leq f-e$. Hence, $\hat{\pi}(\beta, \beta) > \hat{\pi}(\alpha, \beta)$, since $2(b-e) > c$ and, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}, \{a^\beta\}\}$.

If $2(b-e) \leq c \leq \min\{2b, f-e\}$, $\hat{\pi}(\beta, \beta) = (2b-c)/2$ and $\hat{\pi}(\alpha, \beta) = e$. In this case $\hat{\pi}(\beta, \beta) \leq \hat{\pi}(\alpha, \beta)$, since $2(b-e) - c \leq 0$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$.

If $2b < c < f-e$, $\hat{\pi}(\beta, \beta) = 0 < e = \hat{\pi}(\alpha, \beta)$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$.

If $f-e < c < 2b$, $\hat{\pi}(\beta, \beta) = (2b-c)/2 \leq (f+e-c)/2 = \hat{\pi}(\alpha, \beta)$, since $e+f \geq 2b$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$.

If $\max\{2b, f-e\} \leq c < e+f$, $\hat{\pi}(\beta, \beta) = 0 < (f+e-c)/2 = \hat{\pi}(\alpha, \beta)$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$.

If $e+f \leq c < 2d$, $\hat{\pi}(\beta,\beta) = \hat{\pi}(\alpha,\beta) = 0$. Hence, by Lemma 3, $\mathcal{L} = \{\{a^\alpha\}\}$. $\square$

Proof of Theorem 2

Case (I): n odd

I.i) $\frac{c}{2} < r$, *i.e.*, $c < b-e-(d-f)$. We claim that $A^s = \{a^\beta\}$. Note that $b-e-(d-f) < \min\{f-e, 2(b-e)\}$. Since $c < b-e-(d-f) < \min\{f-e, 2(b-e)\}$, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$, $\hat{\pi}(\alpha,\beta) = e$ and $\hat{\pi}(\beta,\alpha) = f-c$. Thus, $c < b-e-(d-f)$ implies $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) > \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$. Hence, by Lemma 5, the claim follows.

I.ii) $\frac{c}{2} > r$, *i.e.*, $c > b-e-(d-f)$. We claim that $A^s = \{a^\alpha\}$. We differentiate six cases:

I.ii.1) $2b > e+f$ and $b-e-(d-f) < c \leq f-e$. Since $c \leq f-e$, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$, $\hat{\pi}(\alpha,\beta) = e$ and $\hat{\pi}(\beta,\alpha) = f-c$. Thus, $c > b-e-(d-f)$ implies $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) < \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$. Hence, by Lemma 5, $A^s = \{a^\alpha\}$.

I.ii.2) $2b > e+f$ and $f-e < c < e+f$. Then, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$ and $\hat{\pi}(\alpha,\beta) = \hat{\pi}(\beta,\alpha) = (f+e-c)/2$. Hence $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) < \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$ and, by Lemma 5, $A^s = \{a^\alpha\}$.

I.ii.3) $2b > e+f$ and $e+f \leq c < 2b$. Then, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$ and $\hat{\pi}(\alpha,\beta) = \hat{\pi}(\beta,\alpha) = 0$. Hence $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) < \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$ and, by Lemma 5, $A^s = \{a^\alpha\}$.

I.ii.4) $2b > e+f$ and $c \geq 2b$. Then, by Proposition 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$.

I.ii.5) $e+f \geq 2b$ and $b-e-(d-f) < c \leq 2(b-e)$. Since $c \leq 2(b-e) \leq f-e$, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$, $\hat{\pi}(\alpha,\beta) = e$ and $\hat{\pi}(\beta,\alpha) = f-c$. Thus, $c > b-e-(d-f)$ implies $\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta) < \hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha)$. Hence, by Lemma 5, $A^s = \{a^\alpha\}$.

I.ii.6) $e+f \geq 2b$ and $c > 2(b-e)$. Then, by Proposition 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$. This proves the claim and the first part of the Theorem.

Case (II): n even and higher than $\bar{n} = \frac{2d-(e+f)}{d-b}$.

II.i) $\frac{c}{2} < r - \delta(n)$, *i.e.*, $c < b-e-(d-f) - \frac{b-e+d-f}{n-1}$. We claim that $A^s = \{a^\beta\}$. Since $c < b-e-(d-f) - \frac{b-e+d-f}{n-1} < \min\{f-e, 2(b-e)\}$, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$, $\hat{\pi}(\alpha,\beta) = e$ and $\hat{\pi}(\beta,\alpha) = f-c$. Since $c < b-e-(d-f) - \frac{b-e+d-f}{n-1}$ implies $(\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))(n-2) > (\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))n$, by Lemma 5, the claim follows.

II.ii) $\frac{c}{2} > r + \delta(n)$, *i.e.*, $c > b-e-(d-f) + \frac{b-e+d-f}{n-1}$. We claim that $A^s = \{a^\alpha\}$. Note that, since $b-e-(d-f) + \frac{b-e+d-f}{n-1} < 2(b-e)$ and $n > \frac{2d-(e+f)}{d-b}$, implies $b-e-(d-f) + \frac{b-e+d-f}{n-1} < f-e$. We differentiate six cases.

II.ii.1) $2b > e+f$ and $b-e-(d-f) + \frac{b-e+d-f}{n-1} < c \leq f-e$. Since $c \leq f-e$, $\hat{\pi}(\alpha,\alpha) = (2d-c)/2$, $\hat{\pi}(\beta,\beta) = (2b-c)/2$, $\hat{\pi}(\alpha,\beta) = e$ and $\hat{\pi}(\beta,\alpha) = f-c$. Since $c > b-e-(d-f) + \frac{b-e+d-f}{n-1}$ implies $(\hat{\pi}(\alpha,\alpha) - \hat{\pi}(\beta,\alpha))(n-2) > (\hat{\pi}(\beta,\beta) - \hat{\pi}(\alpha,\beta))n$, by Lemma 5, $A^s = \{a^\alpha\}$.

II.ii.2) $2b > e + f$ and $f - e < c < e + f$. Then $\hat{\pi}(\alpha, \alpha) = (2d - c)/2$, $\hat{\pi}(\beta, \beta) = (2b - c)/2$ and $\hat{\pi}(\alpha, \beta) = \hat{\pi}(\beta, \alpha) = (f + e - c)/2$. Since $n > \frac{2d-(e+f)}{d-b}$ implies $(\hat{\pi}(\alpha, \alpha) - \hat{\pi}(\beta, \alpha))(n - 2) > (\hat{\pi}(\beta, \beta) - \hat{\pi}(\alpha, \beta))n$, by Lemma 5, $A^s = \{a^\alpha\}$.

II.ii.3) $2b > e + f$ and $e + f \leq c < 2b$. Then $\hat{\pi}(\alpha, \alpha) = (2d - c)/2$, $\hat{\pi}(\beta, \beta) = (2b - c)/2$ and $\hat{\pi}(\alpha, \beta) = \hat{\pi}(\beta, \alpha) = 0$. Note that, in this case, $n > \frac{2d-(e+f)}{d-b}$ implies $n > \frac{2d-c}{d-b}$, which in turn implies $(\hat{\pi}(\alpha, \alpha) - \hat{\pi}(\beta, \alpha))(n - 2) > (\hat{\pi}(\beta, \beta) - \hat{\pi}(\alpha, \beta))n$. Hence, by Lemma 5, $A^s = \{a^\alpha\}$.

II.ii.4) $2b > e + f$ and $c \geq 2b$. Then, by Proposition 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$.

II.ii.5) $e + f \geq 2b$ and $b - e - (d - f) + \frac{b-e+d-f}{n-1} < c \leq 2(b - e)$. Since $c \leq 2(b - e) \leq f - e$, $\hat{\pi}(\alpha, \alpha) = (2d - c)/2$, $\hat{\pi}(\beta, \beta) = (2b - c)/2$, $\hat{\pi}(\alpha, \beta) = e$ and $\hat{\pi}(\beta, \alpha) = f - c$. Since $c > b - e - (d - f) + \frac{b-e+d-f}{n-1}$ implies $(\hat{\pi}(\alpha, \alpha) - \hat{\pi}(\beta, \alpha))(n - 2) > (\hat{\pi}(\beta, \beta) - \hat{\pi}(\alpha, \beta))n$, by Lemma 5, $A^s = \{a^\alpha\}$.

I.ii.6) $e + f \geq 2b$ and $c > 2(b - e)$. Then, by Proposition 3, $A^l = \{a^\alpha\}$. Since $A^s \subseteq A^l$, $A^s = \{a^\alpha\}$. This proves the claim. □

References

[1] Bala V, Goyal S (2000) A non-cooperative model of network formation. Econometrica 68:1181-1231

[2] Battalio R, Samuelson L, van Huyck J (2001) Optimization incentives and coordination failure in laboratory stag hunt games. Econometrica 69:749-764

[3] Bhaskar V, Vega-Redondo F (2004) Migration and the evolution of conventions. Journal of Economic Behavior and Organization 55:397-418

[4] Binmore K, Rubinstein A, Wolinsky A (1986) The Nash bargaining solution in economic modelling. RAND Journal of Economics 17:176-188

[5] Bramoulle Y, López-Pintado D, Goyal S, Vega-Redondo F (2004) Network formation and anti-coordination games. International Journal of Game Theory 33:1-19

[6] Droste E, Gilles R, Johnson K (2000) Evolution of conventions in endogenous social networks. Mimeo. Virginia Tech

[7] Ellison G (1993) Learning, local interaction and coordination. Econometrica 61:1047-1071

[8] Ely J (2002) Local Conventions. Advances in Theoretical Economics 2

[9] Freidlin M, Wentzell A (1984) Random Perturbations of Dynamical Systems. Springer Verlag, New York

[10] Galeotti A, Goyal S, Kamphorst J (2006) Network formation with heterogeneous players. Games and Economic Behavior 54:353-372

[11] Galeotti A, Meléndez-Jiménez M (2004) Exploitation and cooperation in networks. Tinbergen Institute Discussion Paper TI 2004-076/1

[12] Goyal S, Vega-Redondo F (2005) Network formation and social coordination. Games and Economic Behavior 50:178-207

[13] Haag M, Lagunoff R (2006), Social norms, local interaction, and neighborhood planning. International Economic Review 47:265-296

[14] Harsanyi J, Selten R (1988) A General Theory of Equilibrium Selection. MIT Press, Cambridge, MA

[15] Jackson M (2005) A survey of network formation models: stability and efficiency. In: Demange, G and M. Wooders (ed.) Group Formation in Economics: Networks, Clubs and Coalitions, Cambridge University Press

[16] Jackson M, Watts A (2002) On the formation of interaction networks in social coordination games. Games and Economic Behavior 41:265-291

[17] Jackson M, Wolinsky A (1996) A strategic model of economic and social networks. Journal of Economic Theory 71:44-74

[18] Kandori M, Mailath G, Rob R (1993) Learning, mutation and long run equilibria in games. Econometrica 61:29-56

[19] Nash J (1950) The bargaining problem. Econometrica 18:155-162

[20] Nash J (1953) Two person cooperative games. Econometrica 21:128-140

[21] Peski M (2003) State-independent evolution on networks. Northwestern University, mimeo

[22] Peski M (2004) Small group coordination. Northwestern University, mimeo

[23] Slikker M, van den Nouweland A (2000) Network formation models with costs for establishing links. Review of Economic Design 5:333-362

[24] Vega-Redondo F (2006) Building up social capital in a changing world. Journal of Economic Dynamics and Control, 30:2305-2338

[25] Watts A (2001) A dynamic model of network formation. Games and Economic Behavior 34:331-341

[26] Young HP (1993) The evolution of conventions. Econometrica 61:57-84