TI 2000-093/1
Tinbergen Institute Discussion Paper

Learning, Network Formation and Coordination

Sanjeev Goyal
Fernando Vega-Redondo

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at
http://www.tinbergen.nl

Learning, Network Formation and Coordination

Sanjeev Goyal* Fernando Vega-Redondo†

First Version: November 1999.
This version: October 2000.

Abstract

In many economic and social contexts, individual players choose their partners and also decide on a mode of behavior in interactions with these partners. This paper develops a simple model to examine the interaction between partner choice and individual behavior in games of coordination. An important ingredient of our approach is the way we model partner choice: we suppose that a player can establish ties with other players by investing in costly pair-wise links.

We show that individual efforts to balance the costs and benefits of links sharply restrict the range of stable interaction architectures; equilibrium networks are either complete or have the star architecture. Moreover, the process of network formation has powerful effects on individual behavior: if costs of forming links are low then players coordinate on the risk-dominant action, while if costs of forming links are high then they coordinate on the efficient action.

*Econometric Institute, Erasmus University, 3000 DR, Rotterdam, The Netherlands. E-mail address: goyal@few.eur.nl

†Departamento de Fundamentos del Análisis Económico and Instituto Valenciano de Investigaciones Económicas, Universidad de Alicante, 03071 Alicante, Spain & Departament de Economía i Empresa, Universitat Pompeu Fabra, 08005 Barcelona, Spain. E-mail address: vega@merlin.fae.ua.es.

We are deeply indebted to Drew Fudenberg and two anonymous referees for helpful comments. In particular, one of the latter provided suggestions that helped us strengthen, and simplify the proof of, Theorem 3.1. We would also like to thank Bhaskar Dutta, Matt Jackson, Massimo Morelli, Ben Polak, Karl Schlag, Mark Stegeman, Steve Tadelis, and participants in presentations at Alicante, Amsterdam, Center(Tilburg), ISI (Delhi), Stanford, Yale, Econometric Society World Congress (Seattle), and Game Theory World Congress (Bilbao). Vega-Redondo acknowledges support by the Spanish Ministry of Education, CICYT Project no. 970131.

1 Introduction

In recent years, several authors have examined the role of interaction structure – different terms like network structure, neighborhood influences, and peer group pressures, have been used – in explaining a wide range of social and economic phenomena. This includes work on social learning and adoption of new technologies, evolution of conventions, collective action, labor markets, and financial fragility.[1] The research suggests that the structure of interaction can be decisive in determining the nature of outcomes. This leads us to examine the reasonableness/robustness of different structures and is the primary motivation for developing a model in which the evolution of the interaction structure is itself an object of study.

We propose a general approach to study this question. We suppose that individual entities can undertake a transaction only if they are 'linked'. This link may refer to a social or a business relationship, or it may refer simply to awareness of the others. We take the view that links are costly, in the sense that it takes effort and resources to create and maintain them. This leads us to study the incentives of individuals to form links and the implications of this link formation for aggregate outcomes.

In the present paper, we apply this approach to a particular problem: the influence of link formation on individual behavior in games of coordination.[2] There is a group of players, who have the opportunity to play a 2×2 coordination game with each other. Two players can only play with one another if they are 'linked' to each other. These links can be made on individual initiative but are costly to form. So each player prefers that others incur the cost and form links with him. The payoffs of the coordination game are assumed positive and individuals care about aggregate payoffs. Therefore, they always accept any link supported (i.e. paid) by some other player. The link decisions of different players define a network of social interaction. In addition to the choice of links, each player has to choose an action that she will use in all the games that she will engage in. We are interested in the nature of networks that emerge and the effects of link formation on social coordination. In our setting, links as well as actions in the coordination game are

[1] See e.g., Allen and Gale (1998), Bala and Goyal (1998), Chwe (1996), Coleman (1966), Ellison and Fudenberg (1993), Ellison (1993), Granovetter (1974), Haag and Lagunoff (1999), and Morris (2000), among others.

[2] Many games of interest have multiple equilibria. The study of equilibrium selection (which manifests itself most sharply in coordination games) therefore occupies a central place in game theory. We discuss the contribution of our paper to this research in greater detail below.

chosen by individuals on an independent basis. This allows us to study the social process as a population-wide non-cooperative game. Moreover, allowing for links to be formed on an independent basis allows us to distinguish between active and passive links.[3] This distinction plays an important role in our analysis.

We start with a consideration of the situation in which two players can only play a game if they have a direct link between them. We find that a variety of networks – including the complete network, the empty network and partially connected networks – can be supported in equilibria of the static game. Moreover, the society can coordinate on different actions and conformism as well as diversity with regard to actions of individuals is possible in Nash equilibrium. This multiplicity motivates an examination of the dynamic stability of different outcomes.

We develop a dynamic model in which, at regular intervals, individuals choose links and actions to maximize their payoffs. Occasionally they make errors or experiment. Our interest is in the nature of long run outcomes, when the probability of these errors is small. We find that the stochastic dynamics generate clear-cut predictions both concerning the architecture of networks as well as regarding the nature of social coordination.

In particular, we show that the complete network is the unique stochastically stable architecture (except for the case where costs of link formation are very high and the empty network results).[4] Figure 1a gives an example of a complete network in a society with 4 players. This result shows that partially connected networks are not stable. We also find that, if players are at all connected, they always coordinate in the long run on the same action, i.e. social conformism obtains. However, the nature of coordination depends on the costs of link formation. If the costs of link formation are low, players coordinate on the risk-dominant action, while for high costs of link formation they coordinate on the efficient action (Theorem 3.1). Thus our analysis reveals that, even though the eventual network is the same in all cases of interest, the process of network formation itself has serious implications for the nature of social coordination.

[3] Consider two players i and j and suppose player i forms a (costly) link with player j. Such a link is an active link for player i while it is a passive link for player j.

[4] In a complete network, every pair of players is directly linked, while in a empty network there are no links.

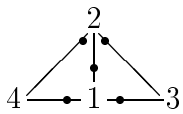

Figure 1a
Complete Network

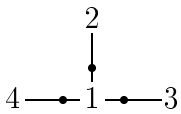

Figure 1b
Center-sponsored Star

The above result says that equilibrium networks are complete or empty. In practice, a variety of factors will lead to incomplete networks; these include capacity constraints on the time and budget of individual players, increasing marginal costs to forming links, and it is also likely that indirect connections will facilitate transactions making complete networks unnecessary. In the present paper we explore the latter possibility, i.e. we investigate the role of *indirect linkages* between players that facilitate transactions. As before, our focus is on the architecture of stable networks and the influence of link formation on the behavior of players in the games with linked players.

Specifically, we consider a model in which two players can play a game if they are directly or indirectly linked with each other.[5] In this setting, we find that the center-sponsored star is the unique stochastically stable architecture. This is a network in which a single player forms a link with every other player and pays for all the links. Figure 1b provides an example of this architecture for a society with 4 players. We also find that there exists a critical cost of forming links, such that, for costs below this level, players coordinate on the risk-dominant action, while for costs above this level, they coordinate on the efficient action (Theorem 4.1).

To summarize: we examine a model in which players choose their partners – by forming costly pairwise links – and choose as well an action in the coordination games played with these partners. Individual efforts to balance the costs and benefits of forming links generate, over time, simple network architectures – the complete network in one case and the star in the other. However, the nature of individual behavior in the coordination games is sensitive to the cost of forming links and this relationship is robust across the settings we examine: players choose the risk-dominant action for low costs of forming links but the

[5]More precisely, two players can play a game with each other if there is a path between them in the social network.

efficient action for high costs. The proofs of our theorems on stochastically stable states use the same general approach. We now sketch this approach briefly.

Suppose that the cost of forming links is such that both types of coordination outcomes, the efficient one as well as the inefficient (risk-dominant) one can be sustained in a social equilibrium. We study the stochastic stability of the two kinds of outcomes: a complete network with efficient coordination and a complete network with inefficient (but risk-dominant) coordination. Roughly speaking, we need to assess the minimum number of "mutations" required to exit from each state. Suppose that we are in the state where everyone is choosing the inefficient action. We assess the minimum number of mutations needed to exit from this state as follows. Given a particular network structure a certain minimum of players must be choosing the efficient action for a player to prefer to play this action. The *first* step is to find a lower bound on the number of such players needed, across the set of all possible networks. This step also derives the network architecture that facilitates the transition. Then, the *second* step shows that the lower bound on mutations computed in the first step is indeed *sufficient* for transition from the particular networks that obtain in equilibrium.

We now develop some intuition for the nature of the network that yields this minimum number of mutations. Here for the sake of concreteness, we focus on the case of direct links. In our model, links are one-sided and this makes them a public good. An action α is particularly attractive for player i when every player choosing α forms a link with i, while no player choosing the alternative action β forms any links with i. In such a situation, if player i were to choose β then she would have to form a link with every player choosing the same action, while if she chooses action α then she can count on 'free-riding' on the links that the others have created.

Number the players from 1 to n. Suppose that a player i forms a link with every other player with a higher index. This generates a complete network. Now suppose that the first k players have their strategies 'mutate', and they all switch from the inefficient action (β) to the efficient action (α). Now consider the situation of player $k+1$. This player is exactly in the situation described above. If the costs of forming links are very small, then player $k+1$ will choose to connect with everyone, irrespective of the choice of actions. Thus the free-riding aspect is relatively unimportant; the network is complete for all practical purposes and standard risk-dominance considerations prevail if cost of forming links is

small. Next suppose that costs are relatively high in the sense that player $k+1$ forms a link with players $k+2,..,n$ *only* if she chooses action β, i.e. the costs of forming links are higher than the miscoordination payoff. In this case, the structure of the network becomes important and we show that, even with relatively few mutants (i.e. the first players k choosing the efficient action), the 'passive' links player $k+1$ enjoys to them makes it attractive for this player (and then all others with higher index) to switch to action α. This in turn leads to the efficient coordination outcome being stochastically stable when costs of forming links are relatively high.

The above considerations suggest that, even among complete networks, there are very significant differences in the relative mutation-responsiveness of each of them, i.e. there can be substantial differences in the minimum mutation threshold required for a global transition to occur. We observe, in particular, that a highly *asymmetric* pattern of connections should generally enhance the mutation fragility of equilibrium configurations. This is why such asymmetric networks play a crucial role in facilitating those transitions that alter the nature of social coordination, e.g. the transition from an inefficient to an efficient state.[6]

We now place the paper and the results in context. Traditionally, sociologists have held the view that individual actions, and in turn aggregate outcomes, are in large part determined by interaction structure. By contrast, economists have tended to focus on markets, where social ties and the specific features of the interaction structures between agents are typically not important. In recent years, economists have examined in greater detail the role of interaction structure and found that it plays an important role in shaping important economic phenomena (see the references given above, and also Granovetter, 1985). This has led to a study of the processes through which the structure emerges. The present paper is part of this general research program.

Next, we relate the paper to work in economics. The paper contributes to two research areas: network formation games and equilibrium selection/coordination problems. We suppose that individual players can form pair-wise links by incurring some costs, at their

[6]This observation is reminiscent of recent work conducted by Albert, Jeong, and Barabási (2000) on the error and attack tolerance displayed by different network arrangements. Specifically, these authors show that the inhomogenous connectivity of many complex networks (e.g. the World-Wide Web, where some nodes bear many links whereas most others only have a few) makes them rather fragile to targeted attack although very tolerant to unguided error. In our case, where mutation probabilities are conceived as very small, the "attack fragility" is the dominant consideration that lends to inhomogeneous networks their key role in the analysis.

own initiative, i.e. link formation is one-sided. This allows us to model the network formation process as a non-cooperative game. This element of our model is similar to the work of Bala and Goyal (2000), from which we borrow some of the techniques applied in the indirect-link scenario. Related work on network formation includes Dutta, van den Nouweland and Tijs (1995) and Jackson and Wolinsky (1996). This earlier work focuses on the architectural and the welfare properties of strategically stable networks, in a context where the sole concern of players is whom they connect to. In contrast, the primary contribution of the present paper is the presentation of a common framework in which the emergence of interaction networks *and* the behaviour of linked players can be jointly studied, both under direct and indirect links.[7]

In many games of interest, multiple equilibria arise naturally and so the problem of equilibrium selection occupies a central place in the theory of games. In recent years, there has been a considerable amount of research on equilibrium selection/coordination;[8] An important finding of this work is that interaction structure and the mobility of players matters and that, by varying the structure, the rate of change as well as the long run outcome can be significantly altered.[9] It is therefore worthwhile to examine the circumstances under which different network structures emerge. From a theoretical point of view, a natural way to do this is by examining the strategic stability of different interaction structures. This is the route taken in the present paper.

A well known result on equilibrium selection in the learning and evolution literature is that risk-dominance considerations prevail over those of efficiency and, if there is a conflict between these two considerations, an inefficient but risk-dominant equilibrium can be stable in the long run. This finding have been re-examined by several authors and the result has been shown to be sensitive to different assumptions, such as the nature of the strategy

[7] In independent work, Droste, Gilles and Johnson (1999), and Jackson and Watts (1999) have developed a related model which addresses similar concerns. The primary difference between these papers and ours pertains to the model of link formation: the former consider two-sided link formation while we study one-sided link formation. Moreover, we allow for direct as well as indirect connections, while those papers consider only direct connections. These differences have a significant impact on the conclusions. We further discuss these papers in the conclusion.

[8] One strand of this work considers dynamic models. This work includes Blume (1993), Canning (1992), Ellison (1993), Kandori, Mailath and Rob (1993), and Young (1993), among others.
For a consideration of this same equilibrium selection problem from a different ("eductive") perspective, the reader may refer to the work of Harsanyi and Selten (1988) or the more recent paper by Carlson and van Damme (1993).

[9] See, for example, Ellison (1993), Goyal (1996) and Morris (2000), among others.

revision rule, precise modeling of mutation, and mobility of players across locations.[10] The results in our paper are closely related to the work on mobility.

The basic insight of the work on player mobility is that if individuals can separate/insulate themselves easily from those who are playing an inefficient action (e.g., the risk-dominant action), then efficient "enclaves" will be readily formed and eventually attract the "migration" of others (who will therefore turn to playing efficiently). In a rough sense, one may be inclined to establish a parallelism between *easy* mobility and *low* costs of forming links. However, the considerations involved in each case turn out to be very different, as is evident from the sharp contrast between our conclusions (recall the above summary) and those of the mobility literature.

There are two main reasons for this contrast. First, in our case, players do not *indirectly* choose their pattern of interaction with others by moving across a *pre-specified* network of locations (as in the case of player mobility). Rather, they construct *directly* their interaction network (with no exogenous restrictions) by choosing those agents with whom they want to play the game. Second, the cost of link formation (which are paid per link formed) act as a screening device that is truly effective only if it is high enough. In a heuristic sense, we may say that it is precisely the restricted "mobility" these costs induce which helps insulate (and thus protect) the individuals who are choosing the efficient action. If the link-formation costs are too low, the extensive interaction they facilitate may have the unfortunate consequence of rendering risk-dominance considerations decisive.

The rest of this paper is organized as follows. Section 2 describes the framework. Section 3 presents the results for the case of direct links, while Section 4 studies the case where players can play a game if they are either directly or indirectly connected to each other. Section 5 concludes.

2 The Model

2.1 Networks

Let $N = \{1, 2, \ldots, n\}$ be a set of players, where $n \geq 3$. We are interested in modelling a situation where each of these players can choose the subset of other players with whom

[10] See e.g., Bergin and Lipman (1996), Robson and Vega-Redondo (1996), Ely (1996), Mailath, Samuelson and Shaked (1994), Oechssler (1997), or Bhaskar and Vega-Redondo (1998), among others.

to play a fixed bilateral game. Formally, let $g_i = (g_{i1}, \ldots g_{i,i-1}, g_{i,i+1}, \ldots g_{in})$ be the set of links formed by player i. We suppose that $g_{ij} \in \{1, 0\}$, and say that player i forms a link with player j if $g_{ij} = 1$. The set of link options is denoted by $\mathcal{G}_\rangle$. Any player profile of link decisions $g = (g_1, g_2 \ldots g_n)$ defines a directed graph, called a *network*. Abusing notation, the network will also be denoted by g.

Specifically, the network g has the set of players N as its set of *vertices* while its set of arrows, $\Gamma \subset N \times N$, is defined as follows: $\Gamma = \{(i, j) \in N \times N : g_{ij} = 1\}$. Graphically, the link (i, j) may be represented as an edge between i and j, a filled circle lying on the edge near agent i indicating that this agent has formed (or supports) that link. Every link profile $g \in \mathcal{G}$ has a unique representation in this manner. Figure 1 below depicts an example. In it, player 1 has formed links with players 2 and 3, player 3 has formed a link with player 1, while player 2 has formed no links.[11]

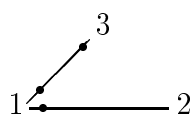

Figure 1

Given a network g, we say that a pair of players i and j are directly linked if at least one of them has established a linked with the other one, i.e. if $\max\{g_{ij}, g_{ji}\} = 1$. To describe the pattern of players' links, it is useful to define a modified version of g, denoted by $\bar{g}$, that is defined as follows: $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}$ for each i and j in N. Note that $\bar{g}_{ij} = \bar{g}_{ji}$ so that the index order is irrelevant. We refer to g_{ij} as an active link for player i and a passive link for player j. We say there is a *path* in g between i and j if either $\bar{g}_{ij} = 1$ or there exist agents $j_1, \ldots, j_m$ distinct from each other and i and j such that $\bar{g}_{i,j_1} = \cdots = \bar{g}_{j_k,j_{k+1}} = \cdots \bar{g}_{j_m,j} = 1$. We write $i \overset{\bar{g}}{\longleftrightarrow} j$ to indicate that there is a path between players i and j in network $\bar{g}$.

Let $N^d(i; g) \equiv \{j \in N : g_{ij} = 1\}$ be the set of players in network g with whom player i has established links, while $\nu^d(i; g) \equiv |N^d(i; g)|$ is its cardinality. Similarly, let $N^d(i; \bar{g}) \equiv \{j \in N : \bar{g}_{ij} = 1\}$ be the set of players in network g with whom player i is *directly linked*, while $\nu^d(i; \bar{g}) \equiv |N^d(i; \bar{g})|$ is the cardinality of this set. Let $N(i; g) \equiv \{j \in N : i \overset{\bar{g}}{\longleftrightarrow} j\}$ be the players with whom player i has a path (is directly or indirectly linked) in a network g; we also define $\nu(i; \bar{g}) \equiv |N(i; \bar{g})|$ to be the cardinality of this set.

[11] Since agents choose strategies independently of each other, two agents may simultaneously initiate a two-way link, as seen in the figure.

A sub-graph $g' \subset g$ is called a *component* of g if for all $i, j \in g'$, $i \neq j$, there exists a path in g' connecting i and j, and for all $i \in g'$ and $j \in g$, $g_{ij} = 1$ implies $g'_{ij} = 1$. A network with only one component is called connected. Given any g, the notation $g + ij$ will denote the network obtained by replacing g_{ij} in network g by 1. Similarly, $g - ij$ will refer to the network obtained by replacing g_{ij} in network g by 0. A connected network g is said to be minimally connected if the network obtained by deleting any single link, $g - ij$, is not connected. A special example of minimally connected network is the center-sponsored star: a network g is called a *center-sponsored star* if there exists some $i \in N$ such that, for all $j \in N \backslash \{i\}$, $g_{ij} = 1$, and for all $j, k \in N \backslash \{i\}$, $j \neq k$, $g_{jk} = 0$.

2.2 Social Game

Individuals located in a social network play a 2×2 symmetric game in strategic form with common action set. The set of partners of player i depends on her location in the network. We shall consider two different models: in the *first* model we will assume that two individuals can play a game if and only if they have a direct link between them. In this case, player i will play a game with all other players in the set $N^d(i; \bar{g})$. In the *second* model a player can play a game with all other players with whom she is directly or indirectly linked. In this case, player i will play a game with the players in the set $N(i; \bar{g})$.

We now describe the bilateral game that is played between any two partners. The set of actions is $A = \{\alpha, \beta\}$. For each pair of actions $a, a' \in A$, the payoff $\pi(a, a')$ earned by a player choosing a when the partner plays a' is given by the following table:

2 / 1	α	β
α	d	e
β	f	b

Table I

We shall assume that the game is one of coordination with two pure strategy equilibria, (α, α) and $\beta, \beta)$. Without loss of generality we will assume that the (α, α) equilibrium is the efficient one. Finally, in order to focus on the interesting case, we will assume that there

is a conflict between efficiency and risk-dominance. These considerations are summarized in the following restrictions on the payoffs.[12]

$$d > f; \quad b > e; \quad d > b; \quad d + e < b + f. \tag{1}$$

An important feature of our approach is that links are *costly*. Specifically, every agent who establishes a link with some other player incurs a cost $c > 0$. Thus, we suppose that the cost of forming each link is independent of the number of links being established and is the same across all players.

Another important feature of our model is that links are *one-sided*. This aspect of the model allows us to use standard solution concepts from non-cooperative game theory in addressing the issue of link formation. We shall assume that the payoffs in the bilateral game are all positive and, therefore, no player has any incentive to refuse links initiated by other players.[13]

Every player i is obliged to choose the same action in the (possibly) several bilateral games that she is engaged in. This assumption is natural in the present context: if players were allowed to choose a different action for every two-person game they are involved in, this would make the behaviour of players in any particular game insensitive to the network structure. Thus, combining the former considerations, the strategy space of a player can be identified with $S_i = \mathcal{G}_\rangle \times \mathcal{A}$, where $\mathcal{G}_\rangle$ is the set of possible link decisions by i and A is the common action space of the underlying bilateral game.[14]

[12]Our results extend in a natural way in case the risk-dominant equilibrium is also efficient, i.e., if $d + e > b + f$. In particular, the network is either complete or a star (depending on the nature of links), while players coordinate on the (α, α) equilibrium, which is risk-dominant as well as efficient, in the long run.

[13]There are different ways in which the assumption of positive payoffs in the coordination game can be relaxed. One route is to dispense with any restriction on payoffs but suppose that, when player i supports a link to player j, the payoff (which may be negative) flows only to i. This formulation may be interpreted as a model of peer groups and fashion, where asymmetric flow of influence seems a natural feature. We have analyzed this model and we find that the relationship between link costs and equilibrium (both on networks and actions) is similar to the one obtained in Theorem 3.1. The only difference pertains to the value of the cut-off value of the cost of forming links identified in the the theorem.
Another possible route to tackle possibly negative payoffs would be to maintain the bi-directional nature of the game but give players the option to refuse the links initiated by others. This would lead to a model with two-sided links which lies outside the scope of the present paper and requires different methods of analysis.

[14]In our formulation, players choose links and actions in the coordination game at the same time. An alternative formulation would be to have players choose links first and then choose actions, contingent on the nature of the network observed. We discuss the timing of moves in the concluding remarks.

We can now present the payoffs of the social game. *First*, we present the payoff for the case where only directly linked players can play with each other. Given the strategies of other players, $s_{-i} = (s_1, \ldots s_{i-1}, s_{i+1}, \ldots s_n)$, the payoff to a player i from playing some strategy $s_i = (g_i, a_i)$ is given by:

$$\Pi_i(s_i, s_{-i}) = \sum_{j \in N^d(i;\bar{g})} \pi(a_i, a_j) - \nu^d(i; g) \cdot c \tag{2}$$

We note that the individual payoffs are aggregated across the games played by him. In much of earlier work, e.g. Kandori, Mailath and Rob (1993) or Ellison (1993), the distinction between average or total payoffs was irrelevant since the size of the neighborhood was given. Here, however, where the number of games an agent plays is endogenous, we want to explicitly account for the influence of the size of the neighborhood and thus choose the aggregate-payoff formulation.

Second, we present the payoffs for the case where two players can play a game if they have a path between them. Given the strategies of other players, $s_{-i} = (s_1, \ldots s_{i-1}, s_{i+1}, \ldots s_n)$, the payoff to a player i from playing some strategy $s_i = (g_i, a_i)$ is then given by:

$$\hat{\Pi}_i(s_i, s_{-i}) = \sum_{j \in N(i;\bar{g})} \pi(a_i, a_j) - \nu^d(i; g) \cdot c \tag{3}$$

These payoff expressions allow us to particularize the standard notion of Nash Equilibrium to each of the two alternative scenarios. Thus, for the model with direct links, a strategy profile $s^* = (s_1^*, \ldots s_n^*)$ is said to be a *Nash equilibrium* if, for all $i \in N$,

$$\Pi_i(s_i^*, s_{-i}^*) \geq \Pi_i(s_i, s_{-i}^*), \quad \forall s_i \in S_i. \tag{4}$$

Similarly, we arrive at the notion of Nash Equilibrium for the indirect-link model by substituting $\hat{\Pi}(.)$ for $\Pi(.)$ in the above expression. On the other hand, a Nash equilibrium in either scenario will be called *strict* if every player gets a strictly higher payoff with her current strategy than she would with any other strategy. The set of Nash equilibria will be denoted by S^* and that of strict Nash equilibria by S^{**}.

2.3 Dynamics

Time is modeled discretely, $t = 1, 2, 3, \ldots$. At each t, the state of the system is given by the strategy profile $s(t) \equiv [(g_i(t), a_i(t))]_{i=1}^n$ specifying the action played, and links established, by each player $i \in N$. At every period t, there is a positive independent probability $p \in (0, 1)$ that any given individual gets a chance to revise her strategy.[15] If she receives this opportunity, we assume that she selects a new strategy

$$s_i(t) \in \arg\max_{s_i \in S_i} \Pi_i(s_i, s_{-i}(t-1)). \tag{5}$$

That is, she selects a best response to what other players chose in the preceding period.[16] If there are several strategies that fulfill (5), then any one of them is taken to be selected with, say, equal probability. This strategy revision process defines a simple Markov chain on $S \equiv S_1 \times \ldots \times S_n$. In our setting, which will be seen to display multiple strict equilibria, there are several absorbing states of the Markov chain.[17] This motivates the examination of the relative robustness of each of them.

To do so, we rely on the approach proposed by Kandori, Mailath and Rob (1993), and Young (1993). We suppose that, occasionally, players make mistakes, experiment, or simply disregard payoff considerations in choosing their strategies. Specifically, we assume that, conditional on receiving a revision opportunity, a player chooses her strategy at random with some small "mutation" probability $\epsilon > 0$. For any $\epsilon > 0$, the process defines a Markov chain that is aperiodic and irreducible and, therefore, has a unique invariant probability distribution. Let us denote this distribution by μ_ϵ. We analyze the form of μ_ϵ as the probability of mistakes becomes very small, i.e. formally, as ϵ converges to zero. Define $\lim_{\epsilon \to 0} \mu_\epsilon = \hat{\mu}$. When a state $s = (s_1, s_2, \ldots, s_n)$ has $\hat{\mu}(s) > 0$, i.e. it is in the support of $\hat{\mu}$, we say that it is *stochastically stable*. Intuitively, this reflects the idea that, even for infinitesimal mutation probability (and independently of initial conditions), this state materializes a *significant* fraction of time in the long run.

[15] This formulation may be interpreted as saying that, with some positive probability, a player dies and is replaced by another player.

[16] We are implicitly assuming that players have complete information the network structure as well as the profile of actions. This assumption simplifies the strategy choice significantly in a setting where a player can potentially play with everyone else in the society.

[17] We note that the set of absorbing states of the Markov chain coincides with the set of strict Nash equilibria of the one-shot game.

3 Direct Links

This section provides an analysis of the model in which two players can undertake a transaction only if they have a direct link between them. We first characterize the Nash equilibrium of the social game. We then provide a complete characterization of the set of stochastically stable social outcomes.

3.1 Equilibrium outcomes

Our first result concerns the nature of networks that arise in equilibria. If costs of link formation are low ($c < e$), then a player has an incentive to link up with other players irrespective of the actions the other players are choosing. On the other hand, when costs are quite high (specifically, $b < c < d$) then everyone who is linked must be choosing the efficient action. This, however, implies that it is attractive to form a link with every other player and we get the complete network again. Thus, for relatively low and high costs, we should expect to see the complete network. In contrast, if costs are at an intermediate level ($f < c < b$), a richer set of configurations is possible. On the one hand, since $c > f (> e)$, the link formation is only worthwhile if other players are choosing the same action. On the other hand, since $c < b (< d)$, coordinating at either of the two equilibria (in the underlying coordination game) is better than not playing the game at all. This allows for networks with two disconnected components in equilibria. The former considerations are reflected by the following result, whose proof is given in Appendix A.

Proposition 3.1 *Suppose (1 and (2) hold. (a) If $c < \min\{f, b\}$, then an equilibrium network is complete. (b) If $f < c < b$, then an equilibrium network is either complete or can be partitioned into two complete components.*[18] *(c) If $b < c < d$, then an equilibrium network is either empty or complete. (d) If $c > d$, then the unique equilibrium network is empty.*

Next, we characterize the Nash equilibria of the static game. First, we introduce some convenient notation. On the one hand, recall that g^e denotes the empty network characterized by $g^e_{ij} = 0$ for all $i, j \in N$ $(i \neq j)$. We shall say that a network g is essential if $g_{ij}g_{ji} = 0$, for every pair of players i and j. Also, let $G^c \equiv \{g : \forall i, j \in N,\ \bar{g}_{ij} = 1,\ g_{ij}g_{ji} = 0\}$ stand for the set of complete and essential networks on the set N. Analogously, for any

[18] Our parameter conditions allow both $f < b$ and $b < f$. If the latter inequality holds, Part (b) of Proposition 3.1 (and also that of Proposition 3.2 below) applies trivially.

given subset $M \subset N$, denote by $G^c(M)$ the set of complete and essential subgraphs on M. Given any state $s \in S$, we shall say that $s = (g, a) \in S^h$ for some $h \in \{\alpha, \beta\}$ if $g \in G^c$ and $a_i = h$ for all $i \in N$. More generally, we shall write $s = (g, a) \in S^{\alpha\beta}$ if there exists a partition of the population into two subgroups, N^α and N^β (one of them possibly empty), and corresponding components of g, g^a and g^β, such that: (i) $g^a \in G^c(N^\alpha)$, $g^\beta \in G^c(N^\beta)$; and (ii) $\forall i \in N^\alpha$, $a_i = \alpha$; $\forall i \in N^\beta$, $a_i = \beta$. With this notation in hand, we may state the following result.

Proposition 3.2 *Suppose (1) and (2) hold. (a) If $c < \min\{f, b\}$, then the set of equilibrium states $S^* = S^\alpha \cup S^\beta$. (b) If $f < c < b$, then $S^\alpha \cup S^\beta \subset S^* \subset S^{\alpha\beta}$, the first inclusion being strict for large enough n. (c) If $b < c < d$, then $S^* = S^\alpha \cup \{(g^e, (\beta, \beta, ..., \beta))\}$. (d) If $c > d$, then $S^* = \{g^e\} \times A^n$.*

Parts (a) and (c) are intuitive; we therefore elaborate on the coexistence equilibria identified in part (b). In these equilibria, there are two unconnected groups, with each group adopting a common action (different in each group). The strategic stability of this configuration rests on the appeal of 'passive' links. A link such as $g_{ij} = 1$ is paid for by player i, but both players i and j derive payoffs from it. In a mixed equilibrium configuration, the links in each group must be, roughly, evenly distributed. This means that all players enjoy some benefits from passive links. In contrast, if a player were to switch actions, then to derive the full benefits of this switch, she would have to form (active) links with everyone in the new group. This lowers the incentives to switch, a consideration which becomes decisive if the number of passive links is large enough (hence the requirement of large n).

The above result indicates that, whenever the cost of links is not excessively high (i.e. not above the maximum payoff attainable in the game), Nash equilibrium conditions allow for a genuine outcome multiplicity. For example, under the parameter configurations allowed in Parts (a) and (c), this multiplicity permits alternative states where either of the two actions is homogeneously chosen by the whole population. Under the conditions of Part (b), the multiplicity allows for a wide range of possible states where neither action homogeneity nor full connectedness necessarily prevails. The model, therefore, raises a fundamental issue of equilibrium selection.

3.2 Dynamics

This section addresses the problem of equilibrium selection by using the techniques of stochastic stability. As a first step in this analysis, we establish convergence of the unperturbed dynamics for the relevant parameter range.

Let $\bar{S}$ denote the set of absorbing states of the unperturbed dynamics. In view of the postulated adjustment process, it follows that there is an one-to-one correspondence between $\bar{S}$ and the class of *strict* Nash equilibria of the social game, i.e. $\bar{S} = S^{**}$. Proposition 3.2 characterizes *all* Nash equilibria of this game. But, clearly, if $c < b$, every Nash equilibrium is strict, while if $b < c < d$, only the Nash equilibria in S^{α} are strict. Since, on the other hand, no strict Nash equilibrium exists if $c > d$, the next result focuses on the case where $c < d$.

Proposition 3.3 *Suppose (1)-(2) hold and $c < d$. Then, starting from any initial strategy configuration, the best response dynamics converges to a strict Nash equilibrium of the social game, with probability one.*

The proof of the above result is given in Appendix A. This result delimits the set of states that can potentially be stochastically stable since, obviously, every such state must be a limit point for the unperturbed dynamics. Let the set of *stochastically stable states* be denoted by $\hat{S} \equiv \{s \in S : \hat{\mu}(s) > 0\}$. The following result summarizes our analysis in the direct-link model.

Theorem 3.1 *Suppose (1) and (2) hold. There exists some $\bar{c} \in (e, b)$ such that if $c < \bar{c}$ then $\hat{S} = S^{\beta}$ while if $\bar{c} < c < d$ then $\hat{S} = S^{\alpha}$, provided n is large enough.*[19] *Finally, if $c > d$ then $\hat{S} = \{g^e\} \times A^n$.*

In order to determine the support of the limit distribution $\hat{\mu}$, we use the well-known graph-theoretic techniques developed by Freidlin and Wentzell (1984) for the analysis of perturbed Markov chains, as applied by the aforementioned authors (Kandori *et al.* and Young) and later simplified by Kandori and Rob (1995). They can be summarized as follows. Fix some state $s \in \bar{S}$. An s-tree is a directed graph on $\bar{S}$ whose root is s and such that there is a *unique* (directed) path joining *any* other $s' \in \bar{S}$ to s. For each arrow $s' \to s''$ in any given s-tree, a "cost" is defined as the minimum number of simultaneous mutations that

[19] The proviso on n is simply required to deal with possible integer problems when studying the number of mutations needed for the various transitions.

are required for the transition from s' to s'' to be feasible through the ensuing operation of the unperturbed dynamics alone. The cost of the tree is obtained by adding up the costs associated with all the arrows of a particular s-tree. The stochastic potential of s is defined as the minimum cost across all s-trees. Then, a state $s \in \bar{S}$ is seen to be stochastically stable if it has the lowest stochastic potential across all $s \in \bar{S}$.

In our framework, individual strategies involve both link-formation and action choices. This richness in the strategy space leads to a corresponding wide variety in the nature of (strict) Nash equilibria of the social game. There are two facets of this variety: (a) we obtain three different types of equilibria in terms of action configuration: S^{α}, S^{β} and $S^{\alpha\beta}$; and (b) there are a large number of strategy profiles that support the complete connectivity prevailing at equilibrium configurations. For example, in a game with 10 players, there are 2^{45} different (link) strategy profiles that can support a complete network. This proliferation of equilibria leads us to develop a simple relationship between the different profiles. In particular, we consider strategy profiles within the sets S^h $(h = \alpha, \beta)$ and show that states in each of these sets can be connected by a chain of single-mutation steps, each such step followed by a suitable operation of the best-response dynamics. To state this result precisely, it is convenient to introduce the metric $d(\cdot)$ on the space of networks that, for each pair of networks g and g', has their respective distance given by $d(g, g') = d(g', g) \equiv \sum_{i,j} |g_{i,j} - g'_{i,j}|/2$. In words, this distance is simply a measure of the number of links that are different across the two networks. With this metric in place, we have:

Lemma 3.1 *For each $s \in S^h$, $h = \alpha, \beta$, there exists an s-tree restricted to S^h such that for all arrows $s' \to s''$ in it, $d(g', g'') = 1$, where g' and g'' are the networks respectively associated to s' and s''.*

The proof of this lemma is given in Appendix A. This lemma implies that, provided $S^h \subset \bar{S}$, the (restricted) tree established by Lemma 3.1 for any $s \in S^h$ involves the minimum possible cost $\left|S^h\right| - 1$. This Lemma also indicates that, in the language of Samuelson (1994), S^{α} (if $c < d$) and also S^{β} (if $c < b$) are *recurrent sets.* This allows each of them to be treated as a single "entity" in the following two complementary senses: (i) if any state in one of these recurrent sets is stochastically stable, so is every other state in this same set; (ii) in evaluating the minimum cost involved in a transition to, or away from, any *given* state in a recurrent set, the sole relevant issue concerns the minimum cost associated to a transition to, or away from, *some* state in that recurrent set. Using (i)-(ii), the analysis of the model

can be greatly simplified. To organize matters, it is useful to consider different ranges for c separately.

Let us start with the case where $0 < c < e$, where the set of absorbing states $\bar{S} = S^\alpha \cup S^\beta$. Since, by Lemma 3.1, the sets S^α and S^β are each recurrent, the crucial point here is to assess what is the minimum mutation cost across all path joining *some* state in S^h to *some* state in $S^{h'}$ for each $h, h' = \alpha, \beta$, $h \neq h'$. Denote these mutation costs by $m^{h,h'}$ and let $\lceil z \rceil$ stand for the smallest integer no smaller than any given $z \in \mathcal{R}_+$. With this notation in place, we state:.

Lemma 3.2 *Suppose that* $0 < c < e$. *Then,*

$$m^{\beta,\alpha} = \left\lceil \frac{b-e}{(d-f)+(b-e)}(n-1) \right\rceil; \quad m^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \right\rceil.$$

Thus, $m^{\beta,\alpha} > m^{\alpha,\beta}$, *for large enough* n.

The proof, given in Appendix A, reflects the standard considerations arising in much of the recent evolutionary theory when the *fixed* pattern of interaction involves every individual of the population playing with all others. Now, if costs are low ($c < e$), such full connectivity is not just assumed but it *endogenously* follows from players' own decisions, both at equilibrium (i.e. when the unperturbed best-response dynamics is at a rest-point) and away from it. In effect, this implies that the same basin-of-attraction considerations that privilege risk-dominance in the received approach also select for it in the present case.

We next examine the case where $e < c < \min\{f, b\}$ where $\bar{S} = S^\alpha \cup S^\beta$. Now, since $c > e$, players who choose action α no longer find it attractive to form links with other players who choose action β. This factor plays a crucial role in the analysis. The following result derives the relative magnitude of the minimum mutation costs.

Lemma 3.3 *Suppose* $e < c < \min\{f, b\}$. *Then,*

$$m^{\beta,\alpha} = \left\lceil \frac{b-c}{(d-f)+(b-c)}(n-1) \right\rceil; \quad m^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \right\rceil.$$

Thus, there is some $\tilde{c}$, $e < \tilde{c} \leq \min\{f, b\}$, *such that if* $c < \tilde{c}$ *then* $m^{\beta,\alpha} - m^{\alpha,\beta} > 0$, *while if* $c > \tilde{c}$ *then* $m^{\beta,\alpha} - m^{\alpha,\beta} < 0$, *for large enough* n.

The methods used to prove this lemma are quite general; we use them in establishing a number of other lemmas, both in the proof of Theorem 3.1 and that of Theorem 4.1. It is therefore useful to explain them in the text.

Proof of Lemma 3.3: Let s^α and s^β be generic states in S^α and S^β, respectively.

Step 1: Consider transitions from state s^β to state s^α and let k be the number of mutations triggering it. If this transition is to take place after those many mutations, there must be some player currently choosing β (i.e. who has *not* mutated) that may then voluntarily switch to α. Denote by q^h the number of active links this player *chooses* to support to players choosing h ($h = \alpha, \beta$) and let r^h stand for the number of passive links she receives from players choosing h ($h = \alpha, \beta$). If she chooses α, her payoff is given by:

$$\pi_\alpha = r^\alpha d + r^\beta e + q^\alpha (d - c), \tag{6}$$

where we implicitly use the fact that q^β must equal zero – since $c > e$, an agent who switches to α will not find it worthwhile to support any link to players choosing β. On the other hand, if the agent in question were to continue adopting β, her payoff would be equal to:

$$\pi_\beta = \hat{r}^\alpha f + \hat{r}^\beta b + \hat{q}^\alpha (f - c) + \hat{q}^\beta (b - c), \tag{7}$$

where $\hat{q}^h$ and $\hat{r}^h$ are interpreted as the active and passive links that would be chosen by the player if she decided to adopt β. Clearly, we must have $r^h = \hat{r}^h$ for each $h = \alpha, \beta$. Thus, if a switch to α is to take place, it must be that

$$\pi_\alpha - \pi_\beta = (r^\alpha + q^\alpha) d - (r^\alpha + \hat{q}^\alpha) f - r^\beta (b - e) - \hat{q}^\beta (b - c) \geq 0. \tag{8}$$

Note that $r^\alpha + q^\alpha = k$, since $c < d$ and therefore the player who switches to α will want to be linked (either passively or by supporting herself a link) to all other players choosing α, i.e. to the total number k of α-mutants. On the other hand, since $c < \min\{f, b\}$, we must also have that $\hat{r}^\beta + \hat{q}^\beta = n - k - 1$ and $r^\alpha + \hat{q}^\alpha = \hat{r}^\alpha + \hat{q}^\alpha = k$, i.e. the player who chooses β must become linked to all other players, both those choosing β and those choosing α.

We now ask the following question: What is the lowest value of k consistent with (8)? Since $c > e$, the desired payoff advantage of action α will occur for the lowest value of k when $r^\beta = \hat{r}^\beta = 0$ and therefore $\hat{q}^\beta = n - k - 1$. That is, if the desired transition is to take place, the *necessary* condition (8) holds for the *minimum* number of required mutations when the arbitrary agent that must start the transition has *no* passive links to individuals

choosing action β. Recall that $m^{\beta,\alpha}$ stands for the minimum number of mutations required for the transition. Now introducing the above observations in (8), we obtain the following lower bound

$$m^{\beta,\alpha} \geq \frac{b-c}{(d-f)+(b-c)}(n-1) \equiv H. \tag{9}$$

The above expression gives the minimum number of players choosing α that are needed to induce some player to switch to action α across all possible network structures. Next, we argue that this number of mutations is also *sufficient* to induce a transition from some s^β to some s^α. The proof is constructive. The main idea is to consider a particular state s^β where its corresponding (complete) network displays the maximal responsiveness to some suitably chosen mutations. Using the observations on the distribution of active and passive links, this is seen to occur when there are some players who support links to all others – those are, of course, players with a "critical" role whose mutation would be most effective. Specifically, suppose that the network prevailing in s^β has every player $i = 1, 2, ..., n$ support active links to all $j > i$. (This means, for example, that player 1 supports links to every other player whereas player n only has passive links.) Then, recalling that $\lceil z \rceil$ denotes the smallest integer no smaller than z, the most mutation-effective way of inducing the population to switch actions from β to α is precisely by having the players $\ell = 1, 2, ..., \lceil H \rceil$ simultaneously mutate to action α and maintain all their links. Thereafter, a transition to some state s^α will occur if subsequent strategy revision opportunities are appropriately sequenced so that every player with index $j = \lceil H \rceil + 1, ..., n$ is given a revision opportunity in order. This, in effect, shows that the lower bound in (9) is tight and $m^{\beta,\alpha} = \lceil H \rceil$.

Step 2: Consider next the transition from some state s^α to a state s^β and let again k be the number of mutations (now towards β) triggering it. Using arguments from Step 1 above, it is easy to show that $m^{\alpha,\beta}$, must satisfy:

$$m^{\alpha,\beta} \geq \frac{d-f}{(d-f)+(b-e)}(n-1) \equiv H'. \tag{10}$$

Again, we can use previous arguments to show that $\lceil H' \rceil$ is sufficient.

Step 3: Finally, we wish to study the difference $m^{\beta,\alpha} - m^{\alpha,\beta}$ as a function of c. For low c (close to e) and large n, this difference is clearly positive in view of the hypothesis that $b - e > d - f$. Next, to verify that it switches strict sign at most once in the range $c \in (e, \min\{f, b\})$, note that $H - H'$ is strictly declining with respect to c in the interval $(e, \min\{f, b\})$. □

Lemma 3.3 applies both to the case where $b < f$ and that where $b > f$. Suppose first that $b < f$. Then, since $H - H' < 0$ for $c = b$, a direct combination of former considerations leads to the desired conclusion for the parameter range $c \in (e, b]$. We now take up the case $f < b$ and focus on the range $c \in (f, b)$. We first derive the relative magnitude of the minimum mutation costs for $s \in S^h$, where $h = \{\alpha, \beta\}$.

Lemma 3.4 *Suppose $f < c < b$.*

$$\tilde{m}^{\beta,\alpha} = \left\lceil \frac{b-c}{(d-f)+(b-c)}(n-1) \right\rceil ; \quad \tilde{m}^{\alpha,\beta} = \left\lceil \frac{d-c}{(d-c)+(b-e)}(n-1) \right\rceil .$$

Thus there is a threshold $\check{c} \in [f, b)$ such that if $c < \check{c}$ then $m^{\beta,\alpha} - m^{\alpha,\beta} > 0$, while if $c > \tilde{c}$ then $m^{\beta,\alpha} - m^{\alpha,\beta} < 0$, for large enough n.

The arguments needed to establish this result are very similar to those used in the proof of Lemma 3.3; we provide the computations in Appendix A.

The principal complication in case $c \in [f, b)$ is that the set of absorbing states in not restricted to $S^\alpha \cup S^\beta$ but will generally include mixed states where the population is segmented into two different action components (cf. Propositions 3.2 and 3.3). Let $m^{h,\alpha\beta}$, for $h = \alpha, \beta$, denote the minimum number of mutations needed to ensure a transition from some $s \in S^h$ to some $s \in S^{\alpha\beta}$. The first point to note is that by the construction used in Lemma 3.3, $m^{\alpha,\alpha\beta} \geq m^{\alpha,\beta}$ and, similarly, $m^{\beta,\alpha\beta} \geq m^{\beta,\alpha}$. This implies that the transition from any state in some S^h towards a mixed equilibrium state in $S^{\alpha\beta}$ is *not* costlier than a transition towards $S^{h'}$ $(h' \neq h)$. Concerning now the converse transitions (i.e. from states in $S^{\alpha\beta}$ to either S^α or S^β), the following lemma indicates that it is relatively "easy" since it involves a suitable chain of *single* mutations..

Lemma 3.5 *Let $f < c < b$ and consider any equilibrium state $s \in S^{\alpha\beta}$ involving two non-degenerate (α and β) components, g^α and g^β, with cardinalities $\mid A(s) \mid > 0$ and $\mid B(s) \mid > 0$, respectively. Then, there is another equilibrium state s' with cardinality for the resulting α component $|A(s')| \geq |A(s)| + 1$ that can be reached from s by a suitable single mutation followed by the best-response dynamics. An identical conclusion applies to some equilibrium state s'' with $\mid A(s'') \mid \leq \mid A(s) \mid -1$.*

The proof of this Lemma is given in Appendix A. We briefly sketch the argument here. Fix some mixed state, and suppose the strategy of some player $i \in A(s)$ mutates as follows: she switches to action β, while everything else remains as before. Now, have all the players in the α group move and suppose that they still wish to keep playing action α. Since $c < f$, their best response is to delete their links with player i. Next, have all the players in group β move; their best response is to form a link with player i. This is true since the original state was an equilibrium, and $c < b$. Finally, have player i choose a best response; since the original state was an equilibrium and $c > f$, her best response is to play action β and delete all links with players in the α group. We have thus increased the number of β players with a single mutation. This argument extends in a natural manner to prove the above result. We now have all the information to complete the proof of the Theorem.

Proof of Theorem 3.1: Consider first the case where $c < b$. If $f > b$, the sets S^α and S^β are the only candidates for stochastic stability and we simply need to compare $m^{\alpha,\beta}$ versus $m^{\beta,\alpha}$. Then, the desired conclusion follows directly from Lemmas 3.2-3.4. The same applies if $f < b$ but $c < f$. Thus, consider the case where $f < c < b$. Then, the states in S^α, S^β, and $S^{\alpha\beta}$ are possible candidates for stochastic stability. Take any state $s \in S^h$ for some $h = \alpha, \beta$. With the help of Lemmas 3.1 and 3.5 we can infer that s-trees for any $s \in S^\alpha$ will have the following minimum cost: $m^{\beta,\alpha} + |S^\alpha| + \left|S^\beta\right| + \left|S^{\alpha\beta}\right| - 2$. For any $s' \in S^\beta$, the situation is symmetric, the minimum cost being equal to $m^{\alpha,\beta} + |S^\alpha| + \left|S^\beta\right| + \left|S^{\alpha\beta}\right| - 2$. Next, concerning any $s \in S^{\alpha\beta}$, we note that the corresponding s-tree would have to display a path joining some state in S^α to s *and* some path joining some state in S^β to s. Thus, the cost of such an s-tree will be at least $m^{\alpha,\beta} + m^{\beta,\alpha} + |S^\alpha| + \left|S^\beta\right| + \left|\tilde{S}^{\alpha\beta}\right| - 1$. This expression is greater than the minimum s-tree costs for $s \in S^h$ $(h = \alpha, \beta)$ since each $m^{h,h'} > 1$ if the population is large. We therefore conclude that a state $s \in S^{\alpha\beta}$ cannot be stochastically stable. Again, therefore, the problem boils down to a comparison of $m^{\alpha,\beta}$ versus $m^{\beta,\alpha}$ which, as before, leads to the desired conclusion.

Next, suppose that $b < c < d$. Then, the key point to observe is that the set of strict Nash equilibria and hence the set of absorbing states is simply $\bar{S} = S^\alpha$. This immediately establishes the result for this case. Finally, similar considerations apply to the case where $c > d$, in which case Propositions 3.2 and 3.3 indicate that $\bar{S} = \{g^e\} \times A^n$. □

In our analysis we have not placed any restrictions on the number of links a player can form and, in equilibrium, the nature of interaction is 'global'. This has the implication that transitions from one strict Nash equilibrium to another require a number of mutations

which is a proportion of the total number of players. As is well known, for large populations this implies that the rate of convergence will be slow. We discuss a model with limited links and its implications for rates of convergence in the conclusion.

4 Indirect Links

In this section, we turn to the context where each player interacts with all players to whom she is joined by a path in the network. We first characterize the Nash equilibria of the social game (Subsection 4.1) and then provide a complete characterization of the set of stochastically stable states (Subsection 4.2).

4.1 Equilibrium outcomes

Our first result derives some basic properties of equilibrium networks and actions.

Proposition 4.1 *Suppose (1) and (3) hold. Then, any equilibrium network is either minimally connected or empty. Furthermore, if the equilibrium network is connected, everyone chooses the same action and social conformism obtains.*

The proofs of all the results in this Subsection (Propositions 4.1-4.3) are given in Appendix B. The above proposition shows that, if any pair of indirectly linked agents play the game, the social disconnectedness and heterogeneity allowed in the direct-link model is no longer possible: in equilibrium, any non-empty equilibrium network must be connected (i.e. define a single component), and every player is to choose the same action. The simple intuition behind this contrast is related to the different appeal of passive links in either model. Recall that, in the setting with direct links, two components displaying different actions could be sustained at equilibrium (if $f < c < b$) because an individual wanting to switch actions and links with the other component must be prepared to forego the benefits of the passive links otherwise springing from her current component. But these lost benefits can only be compensated by bearing the cost of linking to *each* of the members of the other component, a cost that can be prohibitive in those circumstances. Instead, in the setting with indirect links, a player from one component can play with *every* player in the other component by forming a *single* link with someone in that component. Thus, in this case, the considerations arising from passive links cannot overcome, for all players, the large payoff gains to be earned by linking cheaply (i.e. though a single link) with the other component.

The requirement of minimal connectedness allows a wide range of network architectures. This motivates us to examine the requirement of strictness on Nash equilibria, a condition which was obtained "for free" in the direct-link model (when the social network is non-empty). An additional justification for our interest in strict Nash equilibria is that, as explained, all of the rest points of the unperturbed dynamics must correspond to strict Nash equilibria of the social game. The following result shows that 'strictness' is indeed quite effective in our setting.

Proposition 4.2 *Suppose (1) and (3) hold. Then, a strict Nash network is either empty or a center-sponsored star.*

To gain some intuition on the above result, consider any Nash equilibrium network (minimally connected, by virtue of Proposition 4.1) that is not a center-sponsored star. – recall that a center-sponsored star is a network in which a *single* agent supports a link with *every* other player. Then, there has to be a player i who forms a link with some player j, the latter in turn being linked to some third player $k \neq i$. To see that the underlying strategy configuration cannot define a strict Nash equilibrium, simply note that player i can interchange her link with player j for a link with player k and still get the same payoffs. Of course, the only kind of network which is immune to this problem is one in which a single player supports all existing links. A center-sponsored star, in other words, is the only candidate for a strict Nash network.[20]

The above result on strict Nash networks helps us achieve a full characterization of strict Nash equilibria for different values of c. Let G^{cs} stand for the collection of networks that define a center-sponsored star. Furthermore, denote $\tilde{S}^{\alpha} \equiv G^{cs} \times \{\alpha, ..\alpha\}$ and $\tilde{S}^{\beta} \equiv G^{cs} \times \{\beta, ..\beta\}$. Finally, recall that S^{**} represents the set of strict Nash equilibria.

Proposition 4.3 *Suppose (1) and (3) hold. If $0 < c < b$, then $S^{**} = \tilde{S}^{\alpha} \cup \tilde{S}^{\beta}$, while if $b < c < d$, then $S^{**} = \tilde{S}^{\alpha}$. Finally, if $c > d$, then there is no strict Nash equilibrium.*

[20] Our analysis assumes that a direct link is as good as an indirect link. The center-sponsored star remains a strict Nash equilibrium network in more general models of indirect links, where payoffs are lower for games with more distant players. However, in the general setting, it is not be the unique strict Nash network.

The above result parallels, for the present context, the characterization provided by Proposition 3.2 for the direct-link scenario. We note some important differences. First, since all equilibrium networks must be connected (cf. Proposition 4.1), co-existence of the two actions, α and β, is ruled out at any equilibrium configuration. A second interesting difference is that the requirement of strictness amounts here to a genuine refinement criterion (i.e. many Nash equilibria are not strict), even when $c < b$. We do note, however, that for the interesting class of parameters, $c < b$, either actions α or β can be sustained in a strict equilibrium. Next, we address this multiplicity problem.

4.2 Dynamics

In this section we tackle the problem of equilibrium selection, again relying on the techniques of stochastic stability. As a first step in this analysis, we establish convergence of the unperturbed dynamics.

Proposition 4.4 *Suppose (1) and (3) hold. If $c < d$, then starting from any initial strategy configuration the best response dynamics converges to one of the strict equilibrium identified in Proposition 4.3, with probability one. If $c > d$, the best response dynamics converges to the set $\{g^e\} \times A^n$, with probability one.*

The proof of the above result follows from suitable adaptations of arguments used in Theorem 4.1 in Bala and Goyal (2000). These arguments are very long since they involve a consideration of a variety of sub-cases. Here, we simply outline that the four main steps involve establishing that the following transitions have positive probability for the unperturbed dynamics:[21]

- From any given initial state, the process transits to a state in which the network is minimally connected.
- If $c < d$, the process transits from a minimally connected state to one where players are "agglomerated" in the following sense: every two of them are at a distance of no more than 2 (i.e., the maximum number of links between them is two).
- If $c < d$, the process transits from an agglomerated state (as described above) to a center-sponsored star where every agent plays the same action.

[21] The proof of this result is available from the authors upon request.

- If $c > d$, the process transits from any minimally connected state to the empty network.

Proposition 4.4 establishes that the unperturbed dynamics converges, almost surely, to one of the strict Nash equilibria identified in Proposition 4.3. This confines the set of limit points that we need to consider in our analysis of stochastically stability. As before, the set of these stochastically stable states will be denoted by $\hat{S}$.

Theorem 4.1 *Suppose (1) and (3) hold. There exists some $\bar{c} \in (e, b)$ such that if $c < \bar{c}$ then $\hat{S} = \tilde{S}^{\beta}$, while if $\bar{c} < c < d$ then $\hat{S} = \tilde{S}^{\alpha}$, provided n is large enough.*[22] *Finally, if $c > d$ then $\hat{S} = \{g^e\} \times A^n$.*

Thus, in analogy with our analysis for the scenario with direct links, one finds a sharp relationship between the cost of link formation and individual behavior in coordination games. That is, there is a certain threshold $\bar{c} \in (e, b)$ which separates the regions where the inefficient and efficient equilibrium actions are selected. We note, however, an important difference with regard to the architecture of the networks which was already anticipated by our static analysis of the game: if indirect links are allowed, a center-sponsored star is the only robust (here, stochastically stable) architecture that supports the full connectivity required when $c < d$.

Our proof of Theorem 4.1 follows broadly along the same lines as that of Theorem 3.1. First, we note that, in view of Proposition 4.4, only the states in $\tilde{S}^{\alpha}$ (if $c < d$) and also those in S^{β} (if $c < b$) are possible candidates for stochastic stability. Second, we note that the two sets, S^{α} and S^{β}, are recurrent in each case (recall Lemma 3.1 and its ensuing discussion). To verify this claim, assume for concreteness that $c < b$ and consider *any* two states s, $s' \in \tilde{S}^h$, with $h = \alpha, \beta$. We argue that a transition from s to s' can be triggered by a single mutation followed by the operation of the unperturbed dynamics. Specifically, let i and j be the two central players in the center-sponsored stars defined by s and s', respectively. Then, if the central player i mutates at s and removes all her links (still keeping her former action), a transition to s' will materialize provided that, subsequently, player j alone is given a revision opportunity.

[22] The proviso on n is simply required to deal with possible integer problems when studying the number of mutations needed for transition between different states.

As in the direct links setting, the fact that $\tilde{S}^\alpha$ and $\tilde{S}^\beta$ are recurrent states (provided $c < d$ or $c < b$, respectively) simplifies the analysis substantially. Consider first the case where $b < c < d$. Here we know from Proposition 4.4 that $\hat{S} \subset \tilde{S}^\alpha$ and, consequently, since $\tilde{S}^\alpha$ is recurrent, $\hat{S} = \tilde{S}^\alpha$. On the other hand, if $c < b$ (and therefore $c < d$ as well), we have $\hat{S} \subset \tilde{S}^\alpha \cup \tilde{S}^\beta$ and the conclusion hinges upon the *minimum* number of mutations needed to implement the transitions from *some* state $s \in \tilde{S}^h$ $(h = \alpha, \beta)$ to *some* other state $s' \in \tilde{S}^{h'}$ $(h \neq h')$. In analogy with previous notation, denote by $\tilde{m}^{h,h'}$ such a minimum number of mutations. Then, the recurrent set $\tilde{S}^h$ selected (i.e. $\hat{S} = \tilde{S}^h$) is that one for which $\tilde{m}^{h,h'} > \tilde{m}^{h',h}$.

Now, we compute $\tilde{m}^{h,h'}$ $(h, h' = \alpha, \beta;\ h \neq h')$ for different values of c in the interval $(0, b)$. As direct counterparts of the corresponding results established for the direct-link scenario, we have the following results. (Recall that $\lceil z \rceil$ denotes the smallest integer no smaller than z).

Lemma 4.1 *Suppose that* $0 < c < e$. *Then,*

$$\tilde{m}^{\beta,\alpha} = \left\lceil \frac{b-e}{(d-f)+(b-e)}(n-1) \right\rceil; \quad \tilde{m}^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \right\rceil.$$

Lemma 4.2 *Suppose* $e \leq c \leq \min\{f, b\}$. *Then,*

$$\tilde{m}^{\beta,\alpha} = \left\lceil \frac{b-c}{(d-f)+(b-c)}(n-1) \right\rceil; \quad \tilde{m}^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \right\rceil.$$

Lemma 4.3 *Suppose* $f \leq c \leq b$. *Then,*

$$\tilde{m}^{\beta,\alpha} = \left\lceil \frac{b-c}{(d-f)+(b-c)}(n-1) \right\rceil; \quad \tilde{m}^{\alpha,\beta} = \left\lceil \frac{d-c}{(d-c)+(b-e)}(n-1) \right\rceil.$$

The proofs of Lemma 4.1 and Lemma 4.2 are given in Appendix B, while that of Lemma 4.3 is omitted since it is identical to the latter. The general approach is parallel to the proofs of Lemmas 3.2-3.3 above. We first derive the minimum number of mutations needed for transition across all possible network architectures. Then, we show that this minimum number is also sufficient, given the specific star architecture of the equilibrium networks under consideration. The main difference in the argument arises out of the fact that, in making payoff comparisons, individual players are concerned with the size of the component they are connecting to, and not just the payoff from the single game played with the connected player. This alters the individual payoff expressions and necessitates some

additional arguments. The main point, however, is that these alterations in the payoffs do not alter the essential relationship, identified in Theorem 3.1, between costs of forming links and individual behavior in coordination games.

In view of Lemmas 4.1-4.3, the proof of Theorem 4.1 then follows from the observation that, for all $c \in (0, b)$, the minimum number of mutations required to make the transition from $\tilde{S}^h$ to $\tilde{S}^{h'}$ $(h, h' = \alpha, \beta;\ h \neq h')$ in the present indirect-link model is exactly the same as that required to make the transition from S^h to $S^{h'}$ in the former direct-link model. Thus, $\tilde{m}^{h,h'} = m^{h,h'}$ for any c in each of the three parameter ranges, which leads to the desired conclusions by relying on the same arguments as for Theorem 3.1.

We now briefly discuss the similarity in the number of mutations needed for transition in the direct and the indirect links models. For concreteness we consider the case where $e < c < f$ and focus on the transition from s^β to s^α. The main point to note is that a player will link with other players choosing action α irrespective of his own choice of action, while if a player chooses action α then he will only form links with others choosing α and not form links with players who are choosing β. In the direct links model, Lemma 3.3 focuses on the network where the player receives passive links from all the players who have mutated to action α while he has to form active links with each of the players who are still choosing β. In the indirect links model, Lemma 4.2 considers a network where the center and the other players whose strategies have mutated to α are in a single component, while the players who choose β have become isolated due to the deletion of links by the center. In this case if the best-responding player chooses β then he has to form links with each of the other players choosing β. Thus the number of links to be formed with the players who choose β and hence the mutations required for transition are the same in the direct and indirect links models.

5 Concluding Remarks

We develop a model to study the interaction between partner choice and individual behavior in games of coordination. We start with a setting, in which two players can play a game only if they have a direct link between them. We then consider a setting where two players can play the game if they are directly or indirectly linked. Our analysis establishes that individual incentives to economize on link formation costs lead, in both cases, to simple network architectures. In the first context, the unique equilibrium architecture is the

complete network, while in the second one the unique equilibrium architecture is a star. We show that network formation is intimately related to equilibrium selection in both setups: at low costs of forming links, individuals coordinate on the risk-dominant action, while for high costs of forming links individuals coordinate on the efficient action. Thus in each of these two contexts, while the network architecture remains the same, the nature of coordination varies with the cost of forming links. These results suggest that the process of network formation *per se* has powerful implications for the nature of social coordination.

We now discuss some of the main assumptions underlying the analysis. An important aspect of our model is that link formation is one-sided. From a methodological point of view this formulation has the advantage that it allows us to study the social process of link formation and coordination as a non-cooperative game. Moreover, from a substantive point of view this formulation is interesting since it allows for an explicit consideration of the role of active and passive links. Our analysis suggests that this distinction has implications for the coexistence of conventions as well as for the stochastic stability of different outcomes.

In some settings, it is more natural to think of link formation as a two-sided process, i.e. any link being formed leads both of the players involved to incur some costs. In this case both players should acquiesce in the formation of the link. In independent work, Droste, Gilles and Johnson (1999) and Jackson and Watts (1999) study a two-sided link model with direct links (i.e. only directly connected individuals play the game). There are some differences in the details of the two models but the overall analysis is quite similar; thus, we focus our discussion on the paper by Jackson and Watts. They find, like us, that the equilibrium network is complete. However their results on social coordination are quite different. For instance, they find that, if the costs of link formation are high, all those states where players choose a common action (but either of the two) are stochastically stable. Instead, we have found that, when the costs are high (but below the maximum achievable payoff), the only stochastically stable states involve players choosing the efficient action. This contrast arises out of modeling differences in the link formation process and the timing of moves. Specifically, Jackson and Watts postulate that individuals choose links and actions separately, i.e. players choose links taking actions as given while they choose actions taking the links as a given. By contrast, in our setting, any individual undertaking a revision is allowed to impinge on every dimension of her choice and change *both* her action and her supported links. This suggests that it should be interesting to study the effects of varying levels of flexibility in the two choice dimensions, links and

actions– for example, it seems natural to allow for the possibility that link revision might be more rigid than action change.

Another important assumption in our model concerns the number of links allowed or, relatedly, the shape (e.g. concavity or convexity) of the underlying cost function. In our model, we have imposed no limit on the number of links a player can support and the marginal cost of any additional link has been assumed constant. In general, it seems more plausible that players might be constrained in the number of links they can support due to time and resource constraints. Formally, this can be modeled either by directly imposing a restriction on the number of links (by establishing a fixed upper bound) or indirectly deriving it from a sufficiently convex cost function. We feel that such a bounded links model may be more amenable to weaker assumptions concerning information on the network and the action profiles of players. The limited links assumption is likely have important effects on some aspects of the analysis. In the first place, such a formulation is likely to lead to partially connected (sparse) networks being stable. Building upon insights gleaned from existing evolutionary literature (recall the Introduction), this should in turn have significant implications in at least two respects. On the one hand, it may favor the long-run selection of efficiency by allowing the creation of isolated havens or islands, from which efficient behavior can spread throughout the whole population. On the other hand, it is also likely to speed up convergence to the long-run (stochastically stable) states by dispensing with the need of resorting to a large number of simultaneous mutations to trigger the required transitions.

6 Appendix A

Proof of Proposition 3.1: The proof of part (a) follows directly from the fact that $c < f$ and is omitted. We provide a proof of part (b). In this case $f < c < b$. We first show that $a_i = a_j = a$, if i, j belong to the same component. Suppose not. If $\bar{g}_{ij} = 1$, then it follows that the player forming a link can profitably deviate by deleting the link, since $c > f$. Similar arguments apply if i and j are indirectly connected. We next show that if $i \in g'$ and $j \in g''$, where g' and g'' are two components in an equilibrium network g, then $a_i \neq a_j$. If $a_i = a_j$ then the minimum payoff to i from playing the coordination game with j is b. Since $c < b$, player i gains by forming a link, i.e. choosing $g_{ij} = 1$. Thus g is not an equilibrium network. The final step is to note that since there are only two actions in

the coordination game, there can be at most two distinct components. We note that the completeness of each component follows from the assumption that $c < b$.

We next prove part (c). There are two subcases to consider: $c > \max\{b, f\}$ or $b < c < f$. (Note, of course, that the former subcase is the only one possible if $b > f$.) Suppose first that $c > \max\{b, f\}$, and let g be an equilibrium network which is non-empty but also incomplete. From the above arguments in (b), it follows that if $\bar{g}_{i,j} = 1$, then $a_i = a_j = \alpha$. Moreover, if $a_j = \beta$, then player j can have no links in the network. (These observations follow directly from the hypothesis that $c > \max\{b, f\}$.) However, since g is assumed incomplete, there must exist a pair of agents, i and j, such that $\overline{g}_{ij} = 0$. First, suppose that $a_i = a_j = \alpha$. Then, since $c < d$, it is clearly profitable for either of the two players to deviate and form a link with the other player. Suppose next that $a_i = a_j = \beta$. Then, players i and j can have no links and, furthermore, since g is non-empty, there must be at least two other players $k, l \in N$ such that $a_k = a_l = \alpha$. But then player i can increase her payoff by choosing action α and linking to player k. Finally, consider the case where $a_i \neq a_j$ and let player i choose β. Then, if this player deviates to action α and forms a link with player j she increases her payoff strictly. We have thus shown that $\overline{g}_{ij} = 0$ cannot be part of an equilibrium network. This proves that a non-empty but incomplete network cannot be an equilibrium network in the first sub-case considered.

Consider now the case $b < c < f$ and suppose, for the sake of contradiction, that g is an equilibrium network which is non-empty but incomplete. Since $b < c < d$, it follows directly that not every player chooses action α or β. Moreover, in the mixed configuration, all the players who choose α are directly linked (since $c < d$), there is a link between every pair of players who choose dissimilar actions (since $c < f$), but there are no links between players choosing β (since $b < c$). But then it follows that every player choosing β can increase her payoff by switching to action α. This contradicts the hypothesis that the mixed configuration is an equilibrium. This completes the argument for part (c).

Part (d) is immediate from the hypothesis that $c > d$. □

Proof of Proposition 3.2: We start proving Part (a). In view of Part (a) of Proposition 3.1 and the fact that the underlying game is of a coordination type, the inclusion $S^\alpha \cup S^\beta \subset S^*$ is obvious. To show the converse inclusion, take any profile s such that the sets $A(s) \equiv \{i \in N : a_i = \alpha\}$ and $B(s) \equiv \{j \in N : a_j = \beta\}$ are both non-empty. We claim that such an s cannot be an equilibrium.

Assume, for the sake of contradiction, that such a state s is a Nash equilibrium of the game and denote $u \equiv |A(s)|$, $0 < u < n$. Recall from Proposition 3.1 that every Nash network

in this parameter range is complete. This implies that for any player $i \in A(s)$, we must have:

$$(u-1)d + (n-u)e - \nu^d(i;g) \cdot c \geq (u-1)f + (n-u)b - \nu^d(i;g) \cdot c \tag{11}$$

and for players $j \in B(s)$:

$$(n-u-1)b + uf - \nu^d(j;g) \cdot c \geq (n-u-1)e + ud - \nu^d(j;g) \cdot c. \tag{12}$$

It is easily verified that (11) and (12) are incompatible.

Now, we turn to Part (b). The inclusion $S^\alpha \cup S^\beta \subset S^*$ is trivial. To show that the inclusion $S^* \subset S^{\alpha\beta}$ holds strictly for large enough n, consider a state s where both $A(s)$ and $B(s)$, defined as above, are both non-empty and complete components. Specifically, focus attention on those configurations that are symmetric within each component, so that every player in $A(s)$ supports $\frac{u-1}{2}$ links and every player in $B(s)$ supports $\frac{n-u-1}{2}$ links. (As before, u stands for the cardinality of $A(s)$ and we implicitly assume, for simplicity, that u and $n-u$ are odd numbers.) For this configuration to be a Nash equilibrium, we must have that the players in $A(s)$ satisfy:

$$d(u-1) - \frac{u-1}{2}c \geq f\frac{u-1}{2} + b(n-u) - c(n-u) \tag{13}$$

where we use the fact that, in switching to action β, any player formerly in $A(s)$ will have to support herself all links to players in $B(s)$ and will no longer support any links to other players in $A(s)$ – of course, she still anticipate playing with those players from $A(s)$ who support links with him.

On the other hand, the counterpart condition for players in $B(s)$ is:

$$(n-u-1)b - \frac{n-u-1}{2}c \geq du + e\frac{n-u-1}{2} - cu \tag{14}$$

where, in this case, we rely on considerations for players in $B(s)$ that are analogous to those explained before for players in $A(s)$. Straightforward algebraic manipulations show that (13) is equivalent to:

$$\frac{u}{n} \geq \frac{1}{n}\frac{2d-c-f}{2b+2d-3c-f} + \frac{2(b-c)}{2b+2d-3c-f} \tag{15}$$

and (14) is equivalent to:

$$\frac{u}{n} \leq \frac{1}{n}\frac{c+e-2b}{2b+2d-3c-e} + \frac{2b-c-e}{2b+2d-3c-e}. \tag{16}$$

We now check that, under the present parameter conditions:

$$\frac{2b-c-e}{2b+2d-3c-e} > \frac{2(b-c)}{2b+2d-3c-f}. \tag{17}$$

Denote $Y \equiv 2b-c$, $Z \equiv 2b+2d-3c$, and rewrite the above inequality as follows:

$$\frac{Y-e}{Y-c} > \frac{Z-e}{Z-f} \tag{18}$$

which is weaker than:

$$\frac{Y-e}{Y-f} > \frac{Z-e}{Z-f} \tag{19}$$

since $c > f$. The function $\zeta(z) \equiv \frac{z-e}{z-f}$ is uniformly decreasing in z since $b > f > e$. Therefore, since $Y < Z$, (19) obtains, which implies (18). Hence it follows that, if n is large enough, one can find suitable values of u such that (15) and (16) jointly apply. This completes the proof of Part (b).

We now present the proof for part (c). We know from Proposition 3.1 that the complete and the empty network are the only two possible equilibrium networks. Since $c > b > f > e$, it is immediate that, in the complete network, every player must choose α and this is a Nash equilibrium. Then note that, for the empty network to be an equilibrium, it should be the case that no player has an incentive to form a link. This implies that every player must choose β. On the other hand, it is easy to see that the empty network with everyone choosing β is a Nash equilibrium.

The proof of part (d) follows directly from the hypothesis $c > \max\{d, b, f, e\}$. □

Proof of Proposition 3.3: It is enough to show that, from any given state s^0, there is a finite chain of positive-probability events (bounded above zero, since the number of states is finite) that lead to a rest point of the best response dynamics.

Choose one of the two strategies, say β, and denote by $B(0)$ the set of individuals adopting action β at s^0. Order these individuals in some pre-specified manner and starting with the first one suppose that they are given in turn the option to revise their choices (both concerning strategy and links). If at any given stage τ, the player i in question does not want to change strategies, we set $B(\tau+1) = B(\tau)$ and proceed to the next player if some are still left. If none is left, the first phase of the procedure stops. On the other hand, if the player i considered at stage τ switches from β to α, then we make $B(\tau+1) = B(\tau)\backslash\{i\}$ and, at stage $\tau+1$, re-start the process with the first-ranked individual in $B(\tau+1)$, i.e.

not with the player following i. Clearly, this first phase of the procedure must eventually stop at some finite τ_1.

Then, consider the players choosing strategy α at τ_1 and denote this set by $A(\tau_1) \equiv N \backslash B(\tau_1)$. Proceed as above with a chain of unilateral revision opportunities given to players adopting α in some pre-specified sequence, restarting the process when anyone switches from α to β. Again, the second phase of the procedure ends at some finite τ_2.

By construction, in this second phase, all strategy changes involve an increase in the number of players adopting β, i.e. $B(\tau_2) \supseteq B(\tau_1)$. Thus, if the network links affecting players in $B(\tau_1)$ remain unchanged throughout, it is clear that no player in this set would like to switch to α if given the opportunity at $\tau_2 + 1$. However, in general, their network links will also evolve in this second phase, because individual players in $A(\tau_1)$ may form or delete links with players in $B(\tau_1)$. In principle, this could alter the situation of individual members of $B(\tau_1)$ and provide them with incentives to switch from β to α. It can be shown, however, that this is not the case. To show it formally, consider any given typical individual in $B(\tau_1)$ and denote by $\hat{r}^h$, $h = \alpha, \beta$, the number of links received (but not supported) by this player from players choosing action h. On the other hand, denote $\hat{u} \equiv |A(\tau_1)|$. Then, since the first phase of the procedure stops at τ_1, one must have:

$$\begin{aligned} &\max_{q^\alpha, q^\beta} b(q^\beta + \hat{r}^\beta) + f(q^\alpha + \hat{r}^\alpha) - c(q^\alpha + q^\beta) \\ &\geq \max_{q^\alpha, q^\beta} e(q^\beta + \hat{r}^\beta) + d(q^\alpha + \hat{r}^\alpha) - c(q^\alpha + q^\beta) \end{aligned} \tag{20}$$

for all q^α, q^β such that $0 \leq q^\alpha \leq \hat{u} - \hat{r}^\alpha, \quad 0 \leq q^\beta \leq n - \hat{u} - 1 - \hat{r}^\beta$. Now denote by $\tilde{r}^h$ and $\tilde{u}$ the counterpart of the previous magnitudes ($\hat{r}^h$ and $\hat{u}$) prevailing at τ_2. We now show that $\tilde{u} \leq \hat{u}$, $\tilde{r}^\alpha \leq \hat{r}^\alpha$, and $\tilde{r}^\beta \geq \hat{r}^\beta$. First, we note that $\tilde{u} \leq \hat{u}$ by construction of the process. Next note that if $\tilde{r}^\alpha > \hat{r}^\alpha$ then this implies that some player who chooses action α has formed an additional link with player i in the interval between τ_1 and τ_2. This is only possible if $c < e$. It also implies that player i did not have a link with this player at τ_1. This is only possible if $c > f$, a contradiction. Thus $\tilde{r}^\alpha \leq \hat{r}^\alpha$. Finally note that $\tilde{r}^\beta \geq \hat{r}^\beta$ follows from the fact that the all the players choosing β at τ_1 do not revise their decisions in the interval between τ_1 and τ_2.

Therefore, (20) implies:

$$\begin{aligned} &\max_{q^\alpha, q^\beta} b(q^\beta + \tilde{r}^\beta) + f(q^\alpha + \tilde{r}^\alpha) - c(q^\alpha + q^\beta) \\ &\geq \max_{q^\alpha, q^\beta} e(q^\beta + \tilde{r}^\beta) + d(q^\alpha + \tilde{r}^\alpha) - c(q^\alpha + q^\beta) \end{aligned}$$

for all q^α, q^β such that $0 \leq q^\alpha \leq \tilde{u} - \tilde{r}^\alpha$, $\quad 0 \leq q^\beta \leq n - \tilde{u} - 1 - \tilde{r}^\beta$. This allows us to conclude that the concatenation of the two phases will lead the process to a rest point of the best response dynamics, as desired. □

Proof of Lemma 3.1: The proof is constructive. Let $s \in S^h$, $h = \alpha, \beta$, and order in some arbitrary fashion all other states in $S^h \backslash \{s\}$. Also order in some discretionary manner all pairs $(i, j) \in P \times P$ with $i \neq j$. For the first state in $S^h \backslash \{s\}$, say s_1, proceed in the pre-specified sequence across pairs (i, j) reversing the links of those of them whose links are different from what they are in s. This produces a well-defined path joining s_1 to s, whose constituent states define a set denoted by Q_1. Next, consider the highest ranked state in $S^h \backslash Q_1$, say s_2. Proceed as before, until state s_2 is joined to either state s or a state already included in Q_1. Denote the states included in the corresponding path by Q_2. Clearly, when a stage n is reached such that $S^h \backslash (\cup_{\ell=1}^n Q_\ell) = \emptyset$, the procedure described has fully constructed the desired s-tree restricted to S^h. □

Proof of Lemma 3.2: Let s^α and s^β be generic states in S^α and S^β, respectively. We want to determine the minimum number of mutations needed to transit across a pair of them in either direction.

(1). First, consider a transition from s^β to s^α and let k be the number of mutations triggering it. If this transition is to take place *via* the best-response dynamics after those many mutations, there must be some player currently choosing β (i.e. who has *not* mutated) that may then voluntarily switch to α. As before, denote by q^h the number of active links this player supports to players choosing h ($h = \alpha, \beta$) and let r^h stand for the number of passive links she receives from players choosing h ($h = \alpha, \beta$). The payoff from choosing α for that player is given by:

$$\pi_\alpha = r^\alpha d + r^\beta e + q^\alpha (d - c) + q^\beta (e - c). \tag{21}$$

On the other hand, the payoff to choosing β is given by:

$$\pi_\beta = \hat{r}^\alpha f + \hat{r}^\beta b + \hat{q}^\alpha (f - c) + \hat{q}^\beta (b - c), \tag{22}$$

where $\hat{q}^h$ and $\hat{r}^h$ have the same interpretation of active and passive links as before, now associated to the possibility that the player chooses β. Clearly, we have $q^h = \hat{q}^h$ and $r^h = \hat{r}^h$ for each $h = \alpha, \beta$. Concerning the passive links, this is immediate; for active links, it follows from the fact that, since $c < e$, a player will want to create links to all unconnected players,

independently of what they do. Analogous considerations also ensure that (i). $r^\alpha + q^\alpha = k$ and (ii). $r^\beta + q^\beta = n - k - 1$. Thus, in sum, for a transition from some state in S^β to a state in S^α to be triggered, one must have:

$$\begin{aligned} \pi_\alpha - \pi_\beta &= (r^\alpha + q^\alpha)(d - f) - (r^\beta + q^\beta)(b - e) \\ &= k(d - f) - (n - k - 1)(b - e) \geq 0 \end{aligned}$$

Let $m^{\beta,\alpha}$ stand for the minimum number of mutations which lead to such a transition. The above considerations imply that

$$m^{\beta,\alpha} \geq \frac{b - e}{(d - f) + (b - e)}(n - 1), \tag{23}$$

which gives us the minimum number of mutations that are *necessary* for a transition from *any* state s^β to some s^α. However, denoting by $\lceil z \rceil$ the smallest integer no smaller than z, suppose that the *strategies* of $\lceil \frac{b-e}{(d-f)+(b-e)}(n-1) \rceil$ players undergo a simultaneous mutation from any particular state s^β (i.e. these players maintain their links but switch from β to α). Thereafter, the repeated operation of the best-response dynamics is sufficient to induce a transition to a state s^α. Thus the necessary number of mutations computed above is also sufficient to induce a transition from any s^β to some s^α. That is, the inequality in (23) holds with equality.

(2). Consider on the other hand, the transition s^α to s^β. Using the expressions (21) and (22), we can deduce that the minimum number of mutations $m^{\alpha,\beta}$ needed to transit from some state in S^α to a state in S^β satisfies:

$$m^{\alpha,\beta} \geq \frac{d - f}{(d - f) + (b - e)}(n - 1). \tag{24}$$

As in the first case, this gives us the minimum number of mutations needed for a transition. However, consider any state s^α and suppose that the strategies of $\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \rceil$ players undergo a simultaneous mutation (i.e. they maintain their links but switch from α to β). It again follows that the operation of the best-response dynamics suffices to induce a transition to a state s^β. That is, (24) holds with equality.
To conclude, simply note that, if n is large enough,

$$\left\lceil \frac{b - e}{(d - f) + (b - e)}(n - 1) \right\rceil < \left\lceil \frac{d - f}{(d - f) + (b - e)}(n - 1) \right\rceil$$

since $d - f < b - e$. □

Proof of Lemma 3.4 (Sketch): The proof proceeds in the same way as the proof of Lemma 3.3. We therefore only spell out the main computations.

(1). First, consider transitions from state s^β to state s^α and let k be the number of mutations triggering it. We focus on a player currently choosing β and aim at finding the most favorable (i.e. least mutation-costly) conditions that would induce him to switch to α. Along the lines explained in the proof of Lemma 3.3, this leads to the following lower bound:

$$m^{\beta,\alpha} \geq \frac{b-c}{(d-f)+(b-c)}(n-1) \equiv H, \tag{25}$$

which again can be seen to be tight in the sense that, in fact, $m^{\beta,\alpha} = \langle H \rangle$ – recall that $\lceil z \rceil$ stands for the smallest integer no smaller than z.

(2). Analogous considerations for a transition from state s^α to state s^β leads to the lower bound

$$m^{\alpha,\beta} \geq \frac{d-c}{(b-e)+(d-c)}(n-1) \equiv H', \tag{26}$$

which is also tight, i.e. $m^{\beta,\alpha} = \lceil H' \rceil$.

(3). Finally, to study how the sign of $m^{\beta,\alpha} - m^{\alpha,\beta}$ changes for large n as a function of c, note that

$$H - H' \equiv \Delta(c) = \frac{(b-c)(b-e)-(d-f)(d-c)}{[(d-f)+(b-c)]\,[(b-e)+(d-c)]}(n-1). \tag{27}$$

Observe that the denominator of $\Delta(c)$ is always positive, the numerator is decreasing in c, and is moreover negative at $c = b$. This completes the proof. □

Proof of Lemma 3.5: Fix some $s \in S^{\alpha\beta}$, with the players $A(s)$ and $B(s)$ of the α and β components displaying respective cardinalities $|A(s)| \equiv u > 0$ and $|B(s)| \equiv n - u > 0$, respectively. To address the first part of the Lemma, suppose that a player $i \in B(s)$ experiences a mutation, which has the effect of switching her action from β to α and the deletion of all her links with players in $B(s)$. Now consider the players in the set $B(s)\backslash\{i\}$. There are two possibilities: either all of them wish to retain action β, or there is a player who wishes to switch actions.

In the former case, let all of them move and they will retain their earlier strategy except for one change: they will each delete their link with player i, since $f < c < b$. We now get players in $A(s)$ to move and they all form a link with player i, since $f < c < b < d$. It may be checked that we have reached an equilibrium state s', with $A(s') \geq A(s) + 1$.

Consider now the second possibility. Pick a player $j \in B(s)\backslash\{i\}$, who wishes to switch actions from β to α. It follows that this player will delete all her links with players in $B(s)$ and form links with all players in $A(s)$ (since $e < f < c < b < d$). We then examine the incentives of the players still choosing action β, i.e., players in the set $B(s)\backslash\{i,j\}$. If there are no players who would like to switch actions then we repeat step above and arrive at a new state with a larger α-component. If there are players who wish to switch actions from β to α then we get them to move one at a time. Eventually, we arrive at either a new state $s' \in S^{\alpha\beta}$, or we arrive at a state $s' \in S^{\alpha}$.

In either case, we have shown that starting from a state $s \in S^{\alpha\beta}$, we can move with a single mutation to a state s' such that $A(s') \geq A(s)+1$. Since $s \in S^{\alpha\beta}$ was arbitrary, the proof is complete for the first part. The second conclusion concerning some new equilibrium state s'' with $\mid A(s'') \mid \leq \mid A(s) \mid -1$ is analogous. □

7 Appendix B

Proof of Proposition 4.1: We first show that $a_i = a_j$, if i and j belong to the same component. Suppose not and let $a_i = \alpha$ while $a_j = \beta$. Let there be k players in this component with k_α players choosing action α and k_β $(= k - k_\alpha)$ players choosing action β. The payoff to player i from action α is given by $(k_\alpha - 1)d + k_\beta e - l_i c$, where l_i is the number of links formed by i. If, instead, player i were to choose action β (keeping her links), the payoff would be $(k_\alpha - 1)f + k_\beta b - l_i c$. Since, in equilibrium, player i prefers action α it follows that $(k_\alpha - 1)(d - f) \geq k_\beta(b - e)$. Similar calculations show that since, in equilibrium, player j prefers action β it must be true that $(k_\beta - 1)(b - e) \geq k_\alpha(d - f)$. Given that $d > f$ and $b > e$, this generates a contradiction.

We next show that if an equilibrium network is non-empty then there is only one component, i.e. the network is connected. Fix some equilibrium and suppose that g is the corresponding (non-empty) network, with g' being a non-singleton component in g. Suppose, without loss of generality, that $a_i = \alpha$ for every $i \in g'$. Consider now some $g'' \neq g'$ and assume $a_j = \beta$ for every player $j \in g''$. Let k' and k'' be the cardinality of the two components, g' and g'' respectively. Consider now any particular player $i \in g'$ who forms some links $l_i > 0$ in g' and let $(k'-1)d - l_i c$ be her payoff. Since g is part of an equilibrium, it follows that $(k' - 1)d - l_i c \geq k''b - c$. On the other hand, let $(k'' - 1)b - l_j c$ be the payoff to some player $j \in g''$, where l_j is the number of links that she forms. It follows from the definition of equilibrium that $(k'' - 1)b - l_j c \geq k'd - c$. Bringing the former two

inequalities together we reach a contradiction. The argument is analogous in case $a_j = \alpha$. This proves that a non-empty equilibrium network is connected. The minimality of the equilibrium network follows from the assumption that $c > 0$. □

Proof of Proposition 4.2: From Proposition 4.1, we know that an equilibrium network is either empty or minimally connected. Consider a minimally connected equilibrium network g. Suppose that player i has a link with player j in this network, i.e. $g_{i,j} = 1$. We show that in a strict Nash equilibrium, this implies that player j does not have a link with any other player, i.e., $\overline{g}_{j,k} = 0$ for all $k \neq i$. Suppose there is some player k such that $\overline{g}_{j,k} = 1$. In this case, individual i can simply interchange her link with j for a link with k and get the same payoffs. Thus, the strategy of forming a link with j is *not* a strict best response. Hence g is not a strict Nash network. The above argument also implies that, since g is connected, player i must be linked to every other player directly. The resulting network is therefore a star. Moreover, it also follows that this link must be formed by player i herself. For otherwise, if there is a player k such that $g_{k,i} = 1$, then this player is again indifferent between the link with i and some other agent in the star. This implies that the star must be center-sponsored and completes the proof. □

Proof of Proposition 4.3: First, consider case (a). We know from Proposition 4.1 that every player in a component chooses the same action. We also know that there are only two possible equilibrium architectures, $g \in G^{cs}$ and g^e. Clearly, the empty network cannot be part of a strict Nash equilibrium (see also arguments for part (c) below). Thus the only candidates for strict Nash equilibrium are $s \in G^{cs} \times \{(\alpha, \alpha, ..., \alpha)\}$ or $s \in G^{cs} \times \{(\beta, \beta, ..., \beta)\}$. It is easily checked that any of those are indeed strict Nash equilibria.
Consider case (b) next. Again, the empty network is not sustainable by a strict Nash equilibrium. Then the only candidates are $s \in G^{cs} \times \{(\alpha, \alpha, ..., \alpha)\}$ or $s \in G^{cs} \times \{(\beta, \beta, ..., \beta)\}$. It is immediate to see that none of the latter is sustainable as an equilibrium since $c > b$, which implies that the central player does not have an incentive to form a link with isolated players. Thus the only remaining candidates are the former states, which are easily checked to be strict Nash equilibria.
Finally, consider case (c). If $c > d$, then the center-sponsored star cannot be an equilibrium network. Thus, the only candidate for a strict equilibrium network is the empty one. However, if a network is empty, the choice of actions is irrelevant. This means that there is no strict Nash equilibrium in this case. The proof is complete. □

Proof of Lemma 4.1: Let s^α and s^β be generic states in $\tilde{S}^\alpha$ and $\tilde{S}^\beta$, respectively.

Step 1: First, we focus on the transitions from s^β and s^α. Fix some network g and choose a player $i \in N$. Consider the network $g - g_i$ derived from g by the deletion of all of player i's links. Suppose that this latter network has L components, $C_1, C_2, ..., C_L$, with C_1 corresponding to the component of player i. Furthermore, denote by $x(h)$ the total number of players in C_1 who choose action $h = \alpha, \beta$. Similarly, let $y(h)$ stand for the total number of players in $N \backslash C_1 = \cup_{l=2}^{L} C_l$ who choose action $h = \alpha, \beta$ in s^α. Suppose that player i is given a revision opportunity. With the above notation in hand, we may write her maximum payoff from choosing α as follows:

$$\pi_\alpha = x(\alpha)d + x(\beta)e + y(\alpha)d + y(\beta)e - (L-1)c, \tag{28}$$

where we use the fact that $c < e$ and, therefore, player i must find it optimal to link to all components. On the other hand, the maximum payoff to choosing β is given by:

$$\pi_\beta = x(\alpha)f + x(\beta)b + y(\alpha)f + y(\beta)b - (L-1)c. \tag{29}$$

To initiate a transition towards s^α, we must have that player i prefers action α. This may be written as follows:

$$\pi_\alpha - \pi_\beta = (x(\alpha) + y(\alpha))(d - f) - (x(\beta) + y(\beta))(b - e) > 0. \tag{30}$$

We are interested in a network structure which requires the minimum number of players who are choosing α. Let $x(\alpha) + y(\alpha) = k$ and, therefore, $x(\beta) + y(\beta) = n - k - 1$. From the above expression it follows that the minimum value of k for which (30) holds is insensitive to the particular network structure and only depends on the number of players choosing different actions. From (30)we can also infer that the minimum number $\tilde{m}^{\beta,\alpha}$ of simultaneous mutations required to move from any $s^\beta \in \tilde{S}^\beta$ to some state $s^\alpha \in \tilde{S}^\alpha$ is given by :

$$\tilde{m}^{\beta,\alpha} \geq \frac{b-e}{(d-f)+(b-e)}(n-1). \tag{31}$$

From the above discussion it also follows that the payoff comparisons are insensitive to the precise distribution of active and passive links. This implies also that any the number of mutations identified in (31) is sufficient to trigger the desired transition.

Step 2: Consider, on the other hand, the transition from s^α to s^β. Using again the expressions (28) and (29), we can deduce that the minimum number of mutations required (also sufficient) is given by:

$$\tilde{m}^{\alpha,\beta} \geq \frac{d-f}{(d-f)+(b-e)}(n-1). \tag{32}$$

Combining (31) and (32), the desired conclusion follows. □

Proof of Lemma 4.2: First, we extend former notation. Let g be some arbitrarily given network and $i \in N$ a given player in the population. Again, we focus on the network $g - g_i$ derived from g by the deletion of all of player i's links, and let C_1 be the component of player i in $g - g_i$, denoting by $x(h)$ the number of players who choose action h in C_1. Now, however, it is useful to classify the *remaining* $L-1$ components, $C_2, C_3, ..., C_L$, into different categories depending on the mix of actions they display. Specifically, let C_l^h, $h = \alpha, \beta$, stand for a generic α-component in network $g - g_i$ (i.e. a component in which every player chooses action h) and, similarly, let $C_l^{\alpha\beta}$ refer to component in which some players choose α while others choose β. These components C_l^h are indexed by $l = 1, 2, ..., L^h$, where $h = \alpha, \beta, \alpha\beta$. (Note that $L^\alpha + L^\beta + L^{\alpha\beta} = L - 1$.) Furthermore, for each these components C_l^h, the number of players choosing action h' ($h' = \alpha, \beta$) is denoted by $y_l^h(h')$ – hence, for example, $y_l^\alpha(\beta) = 0$ for all l. Finally, we aggregate across different components and make $y^h(h') \equiv \sum_{l=1,2,...,L^h} y_l^h(h')$ and $y(h') \equiv y^{h'}(h') + y^{\alpha\beta}(h')$

Let s^α and s^β be generic states in $\tilde{S}^\alpha$ and $\tilde{S}^\beta$, respectively.

<u>Step 1:</u> We start with transitions from s^β to s^α. Fix some network g and suppose player i receives an opportunity to revise her strategy. Then note that, since we assume that $c > e$, there exists some number $z \geq 2$ such that $(z-1)e < c \leq ze$ and, therefore, if player i chooses action α, she will not form any links with components $C_l^\beta \neq C_1$ whose cardinality $\left|C_l^\beta\right| < z$. This motivates dividing the set of C_l^β components into two groups, small and large, depending on whether their cardinality is above or below the number z. We index the small C_l^β components from 1 to $\check{L}^\beta$, while the large components are indexed from $\check{L}^\beta + 1$ to $\check{L}^\beta + \hat{L}^\beta (= L^\beta)$. Furthermore, we define $\check{y}^\beta(\beta) = \sum_{l=1,...,\check{L}^\beta} y_l^\beta(\beta)$ and $\hat{y}^\beta(\beta) = \sum_{l=\check{L}^\beta+1,...,L^\beta} y_l^\beta(\beta)$. With this notation in place, the payoff to player i of choosing α may be written as follows:

$$\pi_\alpha = x(\alpha)d + x(\beta)e + [y^\alpha(\alpha) + y^{\alpha\beta}(\alpha)]d + [y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta)]e - [L^\alpha + L^{\alpha\beta} + \hat{L}^\beta]c. \quad (33)$$

On the other hand, the payoff from choosing β is equal to:

$$\pi_\beta = x(\alpha)f + x(\beta)b + [y^\alpha(\alpha) + y^{\alpha\beta}(\alpha)]f + [y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta) + \check{y}^\beta(\beta)]b - [L^\alpha + L^{\alpha\beta} + \hat{L}^\beta + \check{L}^\beta]c. \quad (34)$$

To initiate the transition towards s^α, player i must prefer action α to β, i.e. $\pi_\alpha - \pi_\beta > 0$. Using (33)-(34) and setting $k = x(\alpha) + y^\alpha(\alpha) + y^{\alpha\beta}(\alpha)$, this inequality can be rewritten as follows:

$$k(d-f) - x(\beta)(b-e) - [y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta)](b-e) - [\check{y}^\beta(\beta)b - \check{L}^\beta c] > 0. \tag{35}$$

As before we wish to minimize the value of k, conceived as the number of simultaneous mutations towards action α that perturb the state s^α. This in turn means that we aim at minimizing the value of the negative terms in (35). We begin by noting that, for a fixed value of $\check{y}^\beta(\beta)$, the value of the term $[\check{y}^\beta(\beta)b - \check{L}^\beta c]$ is minimized when $\check{L}^\beta = \check{y}^\beta(\beta)$, i.e. when each of the small components is a singleton. This allows us to rewrite (35) as follows:

$$k(d-f) - x(\beta)(b-e) - [y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta)](b-e) - \check{y}^\beta(\beta)[b-c] > 0. \tag{36}$$

We next note that for any fixed value of $x(\beta) + y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta) + \check{y}^\beta(\beta)$, the value of k is minimized when we set the number $x(\beta) + y^{\alpha\beta}(\beta) + \hat{y}^\beta(\beta) = 0$, i.e. when $\check{y}^\beta(\beta) = n-k-1$. This follows from the fact that $b-e > b-c$. Combining these observations, we find that the minimum number of mutations $\tilde{m}^{\beta,\alpha}$ required for the contemplated transition must satisfy:

$$\tilde{m}^{\beta,\alpha} \geq \frac{b-c}{(d-f)+(b-c)}(n-1). \tag{37}$$

We now show that a number of mutations satisfying the above inequality is also *sufficient*, if those mutations are appropriately chosen. Recall that $s^\beta \in \tilde{S}^\beta$ is a center-sponsored star. Let player n be the center of the star and suppose that the following simultaneous mutations occur. On the one hand, $\tilde{m}^{\beta,\alpha} - 1$ players at the spokes switch their action from β to α. On the other hand, player n's strategy also undergoes a mutation: she switches to α and retains her links with the $\tilde{m}^{\beta,\alpha} - 1$ players who have switched actions but deletes all her links with the remaining $n - \tilde{m}^{\beta,\alpha}$ players (who are still playing action β). This pattern of $\tilde{m}^{\beta,\alpha}$ mutations results in a network where the players choosing action α form a center-sponsored star, while all the players choosing β are rendered as singleton components. If these players are then picked for a revision opportunity, the computations leading to (37) imply that they will all choose action α and become linked to the α-component. Subsequently, by Proposition 4.4, the unperturbed dynamics alone is enough to lead the process a.s. to a center-sponsored star with everyone choosing action α. Thus, in sum, we conclude that $\tilde{m}^{\beta,\alpha}$ mutations satisfying (37) are sufficient for a transition from any $s^\beta \in \tilde{S}^\beta$ to some $s^\alpha \in \tilde{S}^\alpha$.

Step 2: Consider next the transition from s^α to s^β. Here, we would like to induce the particular player i who receives a revision opportunity to choose β. Again, her payoffs π_α and π_β from choosing either action are given by (33) and (34). Thus, the required inequality $\pi_\beta - \pi_\alpha > 0$ can be rewritten as follows:

$$-[x(\alpha)+y^\alpha(\alpha)+y^{\alpha\beta}(\alpha)](d-f)+[x(\beta)+y^{\alpha\beta}(\beta)+\hat{y}^\beta(\beta)](b-e)+[\check{y}^\beta(\beta)b-\check{L}^\beta c] > 0. \quad (38)$$

Let $k = x(\beta)+y^{\alpha\beta}(\beta)+\hat{y}^\beta(\beta)+\check{y}^\beta(\beta)$. We wish to minimize the value of k. This, on the one hand, amounts to the minimization of the first negative term in (38). But since the value of $[x(\alpha) + y^\alpha(\alpha) + y^{\alpha\beta}(\alpha)] = n - k - 1$ is insensitive to the precise links of the α-players, their specific distribution across the different components is irrelevant. We can therefore simplify by setting $x(\alpha) = y^{\alpha\beta}(\alpha) = 0$. Next, we take up the other terms, for which we must identify the "best distribution" of the β-players leading to a minimum k in (38). First, we focus on the number $\check{L}^\beta$ of small β-components. Since only if player i chooses β will she link to any of these components, the *net* payoff gain she would enjoy through each of them by choosing β rather than α is $rb - c$, where r stands for the cardinality of the (small) component in question. On the other hand, if those r players were instead part of a large component, player i would link to them *both* if she plans to play α or β. Consequently, the net gain obtained through them by choosing β rather than α would be $r(b - e)$. Combining both considerations, we find that the *difference* between these net gains (corresponding to the alternative possibilities that the r players under consideration belong to either a small or a large β-component) is $re - c$. Since, by definition, $r \leq z - 1$ and $(z-1)e < c$, we conclude that the latter difference is negative and therefore the desired distribution of β-players involves no small components, i.e. $\check{L}^\beta = 0$. Introducing this fact in (38), the minimum number of mutations $\tilde{m}^{\alpha,\beta}$ can be shown to satisfy:

$$\tilde{m}^{\alpha,\beta} \geq \frac{d-f}{(d-f)+(b-e)}(n-1). \quad (39)$$

Finally, we argue that this number of mutations is also *sufficient* for the transition. Suppose that, starting from s^α (whose associated network is a center-sponsored star), there is a simultaneous mutation in the strategy of $\tilde{m}^{\alpha,\beta}$ players whereby they switch their action from α to β without altering their links. Since, in particular, the central player retains her links, those mutations result in a network where all players are still connected through a center-sponsored star. Thus, if players who are still playing action α are subsequently provided with a revision opportunity, the computations leading to (39) imply that they

will choose action β. Thereafter, by Proposition 4.4, the unperturbed dynamics will lead the system to a center-sponsored star with everyone choosing action β, i.e. a state in $\tilde{S}^{\beta}$, with probability 1. In sum, therefore, we confirm that $\tilde{m}^{\alpha,\beta}$ mutations are sufficient for a transition from any state $s^{\alpha} \in \tilde{S}^{\alpha}$ to some state $s^{\beta} \in \tilde{S}^{\beta}$. □

References

[1] Albert, R., H. Jeong, and A.-L. Barabási (2000): "Error and attack tolerance of complex networks", *Nature* **406**, 378-381.

[2] Allen, F. and D. Gale (2000), Financial Contagion, *Journal of Political Economy*, **108**, 1, 1-33.

[3] Bala, V. and S. Goyal (2000), A Non-Cooperative Model of Network Formation, *Econometrica*, **68**, 5, 1181-1231.

[4] Bala V. and S. Goyal (1998), Learning from Neighbours, *Review of Economic Studies* **65**, 595-621.

[5] Bergin, J. and B. Lipman, (1996), Evolution with State-Dependent Mutations, *Econometrica*, 64, 943-957.

[6] Bhaskar, V. and F. Vega-Redondo (1998), Migration and the evolution of conventions, WP *A Discusión*, Universidad de Alicante.

[7] Blume, L. (1993), Statistical Mechanics of Strategic Interaction, *Games and Economic Behavior* **4**, 387-424.

[8] Canning, D. (1992), Average Learning in Learning Models. *Journal of Economic Theory* **57**, 442-472.

[9] Carlson, H. and and E. van Damme (1993), Global Games and Equilibrium Selection, *Econometrica* **61**, 5, 989-1019.

[10] Chwe, M. (1998), Structure and Strategy in Collective Action, mimeo, University of Chicago.

[11] Coleman, J. (1966), *Medical Innovation: A Diffusion Study*, Second Edition, Bobbs-Merrill, New York.

[12] Droste, E, R. Gilles, and K. Johnson (1999), Endogenous interaction and the evolution of conventions, mimeo, Tilburg University.

[13] Dutta, B., A. van Den Nouweland and S. Tijs (1995), Link formation in cooperative situations, *International Journal of Game Theory*, **27**, 245-256.

[14] Ellison, G. (1993), Learning, Local Interaction, and Coordination. *Econometrica* **61**, 1047–1071.

[15] Ellison, G. and D. Fudenberg (1993), Rules of Thumb for Social Learning, *Journal of Political Economy* **101**, 612-644.

[16] Ely, J. (1996), Local Conventions, mimeo Northwestern University.

[17] Freidlin, M. I. and A. D. Wentzell (1984), *Random Perturbations of Dynamical Systems.* Translation of Freĭdlin and Venttsel′ (1979) by J. Szücs. Springer-Verlag, New York.

[18] Goyal, S. (1996), Interaction Structure and Social Change, *Journal of Institutional and Theoretical Economics*, **152**, 3, 472-495.

[19] Granovetter, M. (1974), *Getting a Job: A Study of Contacts and Careers*, Harvard University Press, Cambridge MA.

[20] Granovetter, M (1985), Economic Action and Social Structure: The Problem of Embeddedness, *American Journal of Sociology* **3**, 481-510.

[21] Haag, M. and R. Lagunoff (1999), Social Norms, Local Interaction, and Neighborhood Planning, *mimeo*, Georgetown University.

[22] Harsanyi, J.C. and R. Selten (1988), *A General Theory of Equilibrium Selection*, Cambridge, Mass., MIT Press.

[23] Jackson, M. and A. Wolinsky (1996), A Strategic Model of Economic and Social Networks, *Journal of Economic Theory* **71**, 1, 44-74.

[24] Jackson, M. and A. Watts (1999), On the formation of interaction networks in social coordination games, mimeo, Caltech.

[25] Kandori, M. and G. J. Mailath and R. Rob (1993), Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* **61**, 29–56.

[26] Kandori, M., and R. Rob, (1995), Evolution of equilibria in the long run: A general theory and applications, *Journal of Economic Theory* **65**, 383-414.

[27] Kirman, A. (1997), The Economy as an Evolving Network, *Journal of Evolutionary Economics* **7**, 339-353.

[28] Mailath, G. Samuelson, L. and Shaked, A., (1994), Evolution and Endogenous Interactions, mimeo., Social Systems Research Institute, University of Wisconsin.

[29] Morris, S. (2000), Contagion, *Review of Economic Studies*, **67**, 1, 57-79.

[30] Oechssler, J (1997), Decentralization and the coordination problem, *Journal of Economic Behavior and Organization* **32**, 119-135.

[31] Robson, A. and F. Vega-Redondo (1996), Efficient Equilibrium Selection in Evolutionary Games with Random Matching, *Journal of Economic Theory* **70**, 65-92.

[32] Samuelson, L. (1994), Stochastic stability in games with alternative best replies, *Journal of Economic Theory* **64**, 35-65.

[33] Young, H. P. (1993), The Evolution of Conventions. *Econometrica* **61**, 57-84.