Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER Journal of Economic Dynamics & Control 30 (2006) 2305–2338
JOURNAL OF Economic Dynamics & Control
www.elsevier.com/locate/jedc

Building up social capital in a changing world

Fernando Vega-Redondo[a,b,*]

[a]*Universidad de Alicante, Instituto Valenciano de Investigaciones Económicas, 03071 Alicante, Spain*
[b]*University of Essex, UK*

Received 26 March 2004; accepted 4 July 2005
Available online 30 September 2005

Abstract

This paper models the dynamic process through which a large society may succeed in building up its 'social capital', i.e. a stable and dense pattern of interaction among its members. Agents play repeatedly an idiosyncratic Prisoner's Dilemma with their neighbors. The social network specifies not only the playing partners but, crucially, also determines how relevant strategic information diffuses or new cooperation opportunities are found. Search plays a key role because the environment is subject to payoff 'volatility' and, therefore, agents must continuously adapt to sustain high payoffs. In this context, we study how the architecture of the social network adapts to changes in the environment. We find that network strategic effects are important in enhancing cooperation. This in turn entails that, endogenously, the architecture of the social network adapts (by displaying more cohesiveness) when the environment deteriorates (say, becomes more volatile). These and other conclusions are first obtained by numerical simulations and then supported by approximate mean-field analysis.

JEL classification: C73; D74; D83; D85

Keywords: Networks; Cooperation; Prisoner's Dilemma; Clustering; Search

*Universidad de Alicante, Instituto Valenciano de Investigaciones Económicas, 03071 Alicante, Spain. Tel.: +34 96 590 3616; fax: +34 96 590 3685.
E-mail address: vega@merlin.fae.ua.es.

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2005.07.004

1. Introduction

The network of agent interaction (the social network, for short) is the backbone of an economic system. Its role is two-fold. On the one hand, of course, it determines how agents *come into contact* to carry out their economic activities. But, complementary to this, the social network also maps how the relevant *information* underlying those activities *flows* among the agents. A proper understanding of many economic phenomena, therefore, requires a good grasp of the reciprocal interplay between network architecture and economic behavior, preferably studied in a dynamic scenario.

In general, the social network must be conceived as an endogenous outcome of agents' decisions, much in the same way as any other economic choice. Networking decisions, however, are particularly interesting in that they display the following two features:

(i) They may be conceived as *investment* decisions, subject to the considerations of cost, expectations, and depreciation.
(ii) They produce unintended large externalities (informational and otherwise) on other agents.

The above considerations suggests conceiving the gradual buildup of the network as an accumulation of social capital. The term 'social capital' has been used in recent times with a variety of different meanings, some of them perhaps too vague or devoid of operational content.[1] Here, I focus on one of the most widely agreed incarnations of this concept. I identify the stock of social capital enjoyed by a certain community with the density and stability of its social network. This, of course, is motivated by the implicit assumption that some dense and stable interaction has positive welfare implications, and should typically be correlated with high overall payoffs. Admittedly, this assumption may not suitable for some applications, as stressed, for example, by Durlauf (1999).

The approach pursued here to study these issues has some similarities with that undertaken by Marsili et al. (2004). In particular, in both cases the model displays the following features:

(1) The accumulation of social capital (i.e. the creation of new links) is conducted through search of profitable opportunities that are to be found by 'moving' along the current network. That is, search is a *network-mediated* process.
(2) The value of existing links is subject to *volatility*, understood as stochastic decay. Therefore, network buildup can be seen as a struggle between search and volatility – a manifestation of the so-called Red Queen Principle (Van Valen

[1]Even though the concept has a longer history, it was James Coleman (see Coleman (1988)) who focused the attention of the sociological, and then economic, literature on the notion of social capital. For a good and recent overview on the use and possible misuse of this concept, see Woolcock (2000), whose discussion mostly focuses on development issues.

(1973)): "... it takes all the running you can do, to keep in the same place" – *Through the Looking Glass...*, Carrol (1872).

What is novel in the present paper is that the value of each link is not exogenous but the outcome of a social game in which strategic incentives are intimately associated to the current architecture of the social network. Specifically, I propose a model where players are involved in a collection of pairwise Infinitely Repeated Prisoner's Dilemmas (IRPD). Every pair of agents directly linked by the social network plays an *idiosyncratic* version of this game (i.e. the stage payoffs typically differ across pairs). Play across all these bilateral games is independent, in the sense that the choices made in each of them (cooperation or defection) are separately adopted at each stage by the players involved. The different games, however, are not *strategically* independent since the behavior of a player in one of the games she plays can be made dependent on what has previously happened in other games.[2] This, of course, presumes that information on past behavior diffuses over the network. I assume, however, that it does *not* diffuse instantaneously but instead 'travels' gradually (one step/link at a time) along the network. Naturally, only when any piece of information arrives to any particular player can the latter's choice be affected by it – say, triggering a punishment to a then-revealed defector.

In this context, it is apparent that the range of strategic behavior that can be supported in the infinitely repeated population game must be crucially dependent on the structure of the underlying network. And reciprocally, of course, the particular network that should prevail – more specifically, which links may be formed or removed – also has to depend on the payoffs that can be earned. To formalize these considerations, I first approach them statically and define the notion of *pairwise-stable network* (PSN), that combines standard ideas from the literature on repeated games with the concept of pairwise stability proposed by the matching and network-formation literatures. Informally, let me simply advance at this point that a PSN is a network in which all of its extant links are stable, in the sense that they induce bilateral cooperation.

The first task undertaken in the paper is to *characterize* those networks that qualify as PSN. I find that rather fine details of the architecture of the network are important to understand pairwise stability. For example, a key factor supporting the stability of a link between two players hinges upon the existence of other 'valuable' neighbors who could punish a deviation without much delay (because they are 'close' to both players). In other words, some measure of network cohesiveness (or generalized clustering) is typically important in supporting network stability.

The insights provided by the PSN (equilibrium) notion are a useful first step in the analysis, but the main aim of the paper is a dynamic analysis. That is, the objective is to shed light on how the interplay between behavior in the IRPD and the forces

[2]This feature of the model is reminiscent of the well-known paper of Bernheim and Whinston (1990), which explore the implications for collusion of multi-market strategic interaction. The relationship with this work is discussed in some detail in Section 6.

underlying link formation jointly shape the social network over time. To address these matters, I postulate a population adjustment process through which agents gradually adapt their behavior to the changing circumstances of their environment. This process, which is taken to proceed on a 'slow' time scale relative to the rate at which the stage game is repeated, consists of the following three components.

(1) *Volatility*. The payoffs of existing links are changed (redrawn afresh) with some independent probability, say $\varepsilon > 0$. This probability – a key parameter of the model – is interpreted as a stylized measure of environmental volatility.
(2) *Search and link creation*. Each player receives, with some independent probability, the opportunity of forming one fresh new link. In that event, she gathers the payoff information stemming from the players she 'knows' (i.e. those in her network component). With some small probability, however, she also obtains information that originated 'far away' (i.e. with someone outside her component).
(3) *Removal of unstable links*. Those links which cannot support cooperation (i.e. do not induce bilateral incentives for it) are eliminated.

The law of motion defined by (1)–(3) is shown to induce a stochastic process which is *ergodic*. Its long-run behavior, therefore, is summarized univocally (i.e. *independently* of initial conditions) by its unique invariant distribution. This invariant distribution is characterized analytically in some particular scenarios – for example, when the support of admissible payoffs is low or when the payoff environment is stable ($\varepsilon = 0$). These cases represent useful benchmarks of comparison, but they are not the most interesting. In general, however, an analytical characterization of the long-run distribution or its direct computation for specific setups seems unfeasible. But, by virtue of ergodicity, there is another route possible: long-run invariant magnitudes of any variable of interest can be 'computed' indirectly through simulations for any given setup. This follows from the fact that, along any simulation path, the empirical averages computed over time must converge almost surely to the theoretical means induced by the invariant distribution.

In the present paper, I rely on such a numerical approach to understand the long-run behavior and trade-offs concerning the following key variables of the model:

- *Network density*, as given by the average degree (or connectivity) of the nodes (agents).
- *Network cohesiveness*, as reflected by the average distance between the neighbors of any given node.
- *Network span*, as embodied by the size of the network components.
- *Payoff performance*, as measured by the average payoff earned per interaction.

The main regularities observed can be succinctly advanced as follows:

(a) The long-run density of the network depends negatively on ε, the extent of payoff volatility, and so happens as well with the average payoff per extant link. Thus, as expected, volatility is found to be detrimental for both the accumulation of social capital and its return.

(b) As volatility rises, the population's (uncoordinated) struggle to preserve their level of connectivity has the *endogenous* (side-)effect of increasing the cohesiveness of the social network, which in turn partially offsets its negative impact.

(c) Whenever the society is able to sustain a dense social network, its architecture displays a wide span – i.e., there is a large component including most of the connected individuals (i.e. those with some social capital).

(d) The harmful effects of volatility are mitigated and arise only gradually by virtue of the enhanced deterrence of opportunistic behavior availed by the social network. If, by way of contrast, every bilateral IRPD game is played independently (i.e. unresponsive to the social network in which it is embedded), those effects are abrupt and much stronger.

The above conclusions underscore the point that a 'stable environment' (i.e. one where agents' payoff conditions do not change too fast) is an important requirement for a successful accumulation of social capital. As volatility increases, the network becomes more cohesive, which can be interpreted as the (uncoordinated) way in which society enhances the 'strategic leverage' of network effects and thus strives to maintain cooperation. Such increased cohesiveness, however, is not achieved at the cost of a narrower network span. The latter is kept as wide as possible, since effective search remains a crucial tool against volatility by allowing for quick adaptation to environmental changes.

As a complement of numerical analysis, the model is also studied analytically by relying on mean-field analytical techniques widely used in statistical physics for the study of complex systems. Very schematically, the aim of mean-field analysis is to formulate a stylized 'model' of the original model where the randomness and micro-detail of the latter is replaced by the expected (and therefore deterministic) motion of a suitably representative construct. Here, I undertake a quite elementary approach along these lines, while referring to Marsili et al. (2004) for an illustration of how a more elaborate approach may yield a wider range of implications. As it turns out, however, the simple mean-field analysis pursued here is in essential accordance with some of the key regularities listed above and is sufficient to shed useful light on the numerical simulations.

The rest of the paper is organized as follows. Section 2 presents the model – first, its static version in Subsection 2.1, then its dynamic counterpart in Subsection 2.2. The study of the model starts in Section 3 with the characterization of pairwise-stable networks. It proceeds with the dynamic analysis in Section 4, which consists of two subsections: Subsection 4.1, that establishes some basic dynamic results (e.g. the ergodicity of the process), and Subsection 4.2 that contains the bulk of our numerical analysis. Next, Section 5 undertakes the mean-field analysis of the model, and compares its conclusions with the simulation results. Finally, in Section 6 the related literature is reviewed, while Section 7 offers some concluding remarks and a number of possible courses for future research.

2. Model

2.1. Static framework

Let $N = \{1, 2, \ldots, n\}$ be a finite population of agents who interact repeatedly in pairs as reflected by the prevailing social network. To fix ideas, we may think of each such bilateral interaction as the development of a certain joint project over time (e.g. the development of a new product) that, given exogenous conditions (affecting, say, market demand or the relevant technology) displays both long-run benefits from cooperation and short-run gains from opportunistic behavior. More precisely, each pair of linked agents, $i, j \in N$, is involved in an infinite repetition of a Prisoner's Dilemma (PD) with *idiosyncratic* payoffs given by the table

$i \backslash j$	C	D
C	ζ_{ij}	$\zeta_{ij} - v$
D	$\zeta_{ij} + v$	0

(1)

where $v > \zeta_{ij} (= \zeta_{ji}) > 0$. As customary, C and D will be labelled as 'Cooperate' and 'Defect,' respectively. Thus, the payoff ζ_{ij} obtained by both players if they jointly cooperate is ij-specific and, in the dynamic ('volatile') environment to be considered later on, it will change over time. For simplicity, the payoff of joint defection is normalized to zero, whereas in case of a unilateral defection the gain v obtained by the defector over ζ_{ij} is made equal to the loss incurred her partner. From a strategic viewpoint, the ratio ζ_{ij}/v is the sole relevant magnitude. Hence one could normalize $v = 1$ and reinterpret the cooperation payoff accordingly. Our whole analysis (cf. in particular Propositions 2 and 1) only depends on this ratio.

The pattern of bilateral interaction (i.e. which pairs of agents actually play) is specified by the *social network*. This network is the result of players' connecting decisions, which are captured by a certain *directed graph* $\vec{g} \subset N \times N$, where the nodes are the players and each directed link $(i,j) \in \vec{g}$ represents the decision of player i to connect to player j. Since, in equilibrium, agents' payoffs must always be non-negative (see Section 3), it is convenient to suppose that every linking decision leads to actual play and therefore the *social network* of inter-agent relations induced by $\vec{g}$ is simply the undirected graph $g \subset N \times N$ defined as follows:

$$\forall i, j \in N, \quad (i,j) \in g \;\Leftrightarrow\; [(i,j) \in \vec{g} \quad \text{or/and} \quad (j,i) \in \vec{g}].$$

For any player i, the set of her neighbors (i.e. the players with whom i interacts) is simply given by:[3]

$$N_i \equiv \{j \in N : (i,j) \in g\}.$$

Usually, the more compact notation ij (or ji) will be used to denote the link between player i and j. Furthermore, we shall write $g - ij$ or $g + ij$ to represent the networks obtained from g by, respectively, adding or removing a link ij.

[3]Note, of course, that if $(i,j) \in g$, it follows that $(j,i) \in g$ as well.

As explained, given the prevailing network g, all pairs of players i and j such that $ij \in g$ are involved in an IRPD with idiosyncratic *stage* payoffs given by their respective ζ_{ij} in (1) and intertemporal flow of these payoffs discounted at a common rate $\delta < 1$. Each of these different IRPD are played 'in parallel,' i.e. one round of each of them is played at every stage. They are *choice independent* in the sense that players' decisions in any one of these games do not restrict the choices that may be taken in others. They need *not* be, however, *strategically independent* since the behavior in any of them may be contingent on the information that is available on what has occurred in other games.

To formulate matters precisely, we need to specify:

- How strategically relevant information about agents' past behavior spreads throughout the network.
- How players make their behavior contingent on that information.

First, concerning the spread of information, the key assumption of the model is that it diffuses through the social network only gradually. Specifically, it is supposed that at the beginning of every stage t, each pair of connected players i and j pool all the strategically relevant information they have acquired so far on the behavior of others (including the partner under consideration). This information thus becomes common knowledge for i and j at stage t and their bilateral behavior can be made contingent on it (see below). In turn, it also increases the stock of information that both players will enjoy at the beginning of the subsequent stage $t + 1$, which will pooled with all of their respective neighbors in an analogous way. Admittedly, the formulation considered is extreme in that it fully abstracts from any strategic considerations that might affect information transmission.[4] I choose it, however, to gain a sharper focus on how incentives impinge on behavior in a network setup. Its key implication is that all relevant information must 'travel' along the network one link at a time and, therefore, the architecture of interaction must have an important bearing on the range of sustainable outcomes.

Second, concerning behavior, players' choices are assumed to be governed by *simple strategy profiles*, in the sense of Abreu (1998). Adapted to our framework, these are profiles of strategies that include, for each link $ij \in g$, the following two items:

(a) An *intended path* of behavior Q^*_{ij}, embodying a sequence of behavior (i.e. action pairs) to be displayed in their interaction over time by the players connected by the link $ij \in g$.
(b) A *punishment path* $\hat{Q}^i_{ij}$, specifying a sequence of behavior to be displayed by the players connected by the link $ij \in g$, right after a unilateral deviation of i becomes *commonly know* by both players, i and j. (Note that, in principle, for every link ij

[4]For a paper that makes the transmission of information an object of strategic choice in a network setup with repeated interaction, see Lippert and Spagnolo (2002). This paper is briefly discussed in Section 6, where I provide a general survey of the related literature.

there could be alternative punishment paths $\hat{Q}_{ij}^{i}$ and $\hat{Q}_{ij}^{j}$, depending on whether it is player i or j who are to be punished.)

Following Abreu (1998), a 'deviation' by player i is taken to occur when either she abandons unilaterally the intended path of play $\hat{Q}_{ik}^{*}$ for some $k \in N_i$ or, alternatively, when she fails to follow the punishment path induced by any other prior deviation (by herself or others). The only adaptation needed in our case concerns the caveat that a punishment path involving players i and j comes into effect only when it becomes common knowledge to both of these players that one of them has deviated. As explained above, the specific delay involved in reaching such a state of bilateral common knowledge depends on the architecture of the social network.

Let

$$\mathcal{Q} \equiv \{(Q_{ij}^{*})_{ij\in g}, (\hat{Q}_{ij}^{i}, \hat{Q}_{ij}^{j})_{ij\in g}\}$$

embody a simple strategy profile (intended and punishment paths) as described above. For short, we shall call $\mathcal{Q}$ a *convention*. As in Abreu (1998), we are interested in subgame-perfect equilibrium *outcomes* that can be implemented though some such convention $\mathcal{Q}$. In this convention, the paths $\hat{Q}_{ij}^{i}$ that are prescribed to ensue after any deviation on the part of player i must belong to a subgame-perfect equilibrium of the continuation game. Abreu called such paths 'penal codes,' highlighting in particular those he labelled *optimal penal codes*, i.e. equilibrium continuation paths that induce a minimum payoff for the deviating party. In our setup, it is clear that penal codes are optimal when, for every $ij \in g$, the corresponding $\hat{Q}_{ij}^{i}$ and $\hat{Q}_{ij}^{j}$ display *joint indefinite* defection. (Thus, we may write simply $\hat{Q}_{ij} \equiv \hat{Q}_{ij}^{i} = \hat{Q}_{ij}^{j}$). Note that this behavior is 'perfect' – i.e. part of a continuation equilibrium – since, once both players come to share the knowledge that a punishment must follow, both have an incentive to pursue unrelenting defection under the anticipation that the partner will do so as well.

The advantage of optimal penal codes is that they span the widest possible range of behavior that can be supported in equilibrium. Thus, in our framework, they provide the best basis to support cooperative behavior in a social network. Specifically, consider any given network g and denote by $Q^{*}(g)$ the path where every player $i \in N$ indefinitely cooperates with each of her neighbors $j \in N_i$. That is,

$$Q^{*}(g) = (Q_{ij}^{*})_{ij\in g}, \quad Q_{ij}^{*} = [(C, C), (C, C), \ldots] \ \forall ij \in g.$$

We are interested in the following question: can optimal penal codes $(\hat{Q}_{ij}^{i})_{ij\in g}$ support the cooperation path $Q^{*}(g)$? Only if this happens for all of the links in g, we shall say that the social network is *pairwise stable*. The implicit idea here is that only if players i and j can ascertain that their link ij will support cooperation in equilibrium will they be interested in maintaining their link.

To formalize matters, consider the following notation. Given a network g and any $ij \in g$, let $Q^{*}(g) \overset{ij}{\setminus} D$ stand for the path derived from $Q^{*}(g)$ when player i unilaterally deviates from Q_{ij}^{*} by defecting with her neighbor j, then following thereafter the remaining part of the convention. This means, in particular, that she follows the induced optimal penal codes when they come into effect but remains cooperative with her alternative neighbors otherwise. So, as is customary in the network literature (cf. Jackson and Wolinsky (1996)), we suppose that players assess the incentives pertaining each link separately and deviations can only affect one of their links at a time.

We can now define the notion of a PSN. But first we need to provide a formal specification of payoffs. Given any path of actions $Q = (Q_{ij})_{ij \in g}$ let $\{\psi_{ij}^{\tau}(Q)\}_{\tau=0}^{\infty}$ represent the flow of stage payoffs earned by player i on the link $ij \in g$ along this path. Then, recalling that $\delta \in (0, 1)$ stands for the common discount factor, the (normalized) discounted payoff induced by that payoff flow is given by $(1-\delta)\sum_{\tau=0}^{\infty} \delta^{\tau}\psi_{ik}^{\tau}(Q)$. Thus, aggregating across all of player i's links, her overall payoffs are as follows:

$$\pi_i(Q) \equiv \sum_{k \in N_i} \left[(1-\delta) \sum_{\tau=0}^{\infty} \delta^{\tau} \psi_{ik}^{\tau}(Q) \right].$$

Definition 1. An undirected graph $g \subset N \times N$ is said be a PSN if for all $ij \in g$,

$$\pi_i(Q^*(g)) \geqslant \pi_i(Q^*(g) \overset{ij}{\setminus} D).$$

Verbally, a PSN is a network where, for every separate link, the two players involved have incentives to use it for cooperation rather than defection. The PSN concept might be criticized as too weak in that it does not allow for a full range of deviations. Specifically, it rules out that a player i may combine a starting deviation for some link ij with a subsequent 'preemptive' deviation on her other links ik, right before the original deviation becomes common knowledge for i and k and thus the penal code $\hat{Q}_{ik}$ starts being carried out. Allowing for this richer set of deviations raises, however, a number of issues. Firstly, it is an indirect way of introducing, in effect, the possibility of multiple-link deviations. And, as explained above, this is at odds with much of the recent network literature where network stability is assessed independently for each link. Secondly, note that the deviations considered would require the use of significantly more complicated strategies, since the deviating player would need to keep track of the precise point in time when each neighbor will learn about the original deviation. Thus, for example, if we were to model the complexity of a strategy through the number of states required to *implement* it through a finite automaton (cf. Neyman (1985), Rubinstein (1986)), arbitrarily complex strategies might be needed to undertake such a deviation if the population is large.[5]

The above comments notwithstanding, it is undoubtedly of interest to explore the implications of a variation of the pairwise-stability notion that requires robustness

[5]Note that, as is customary in the Finite Automata literature in repeated games, this statement concerns the complexity of *implementing* the strategy, not on the complexity of assessing its optimality.

against those richer deviations. Networks that fulfill it will be called strongly PSN. To formalize this alternative stability notion, we adapt previous notation and let $Q^*(g) \overset{ij}{\backslash\backslash} D$ denote the path induced by a deviation on the part of i in the link ij, followed by a preemptive deviation in all other links ik right before the onset of the corresponding (optimal) penal code $\hat{Q}_{ik}$.

Definition 2. An undirected graph $g \subset N \times N$ is said be a *strongly pairwise-stable network* (SPSN) if for all $ij \in g$,

$$\pi_i(Q^*(g)) \geqslant \pi_i(Q^*(g) \overset{ij}{\backslash\backslash} D).$$

In Section 3, we characterize the networks that are pairwise-stable under both the weak and strong notions (Definitions 1 and 2, respectively), finding that the two alternative notions give rise to similar insights. Since the same similar performance carries over as well to network evolution, the dynamic analysis of Section 4 only reports the results arising with the weaker version of pairwise stability.

2.2. Dynamic framework

The equilibrium approach embodied by the PSN concept is essentially static. To enrich the analysis with a dynamic dimension, I propose the following framework.

Time is discrete, with $t = 0, 1, 2, \ldots$ indexing the consecutive time periods. At every t, each agent $i \in N$ supports a certain set of links (her 'active' links), which consist of those that he chose in the past and still maintains. These links belong to the set $\vec{g}_i(t) \subset \{i\} \times N \backslash \{i\}$. As in the static framework, we posit that agent i interacts at t with her neighbors $N_i(t)$ in the *undirected* graph $g(t) = \{jk \in N : jk \in \vec{g}_j(t) \wedge kj \in \vec{g}_k(t)\}$, i.e. active and passive links are equivalent in terms of their implications for play. The sole difference between the active and passive links of a player is that while the former are taken on her own initiative, the passive ones depend on the decisions of others. To bound the state space of the process, it is convenient to postulate that any given player can support at most m active links, with m being a parameter of the model. This may be justified, for example, by supposing that the support of a link requires resources that every agent has in limited supply (e.g. time).

Each of the links $ij \in g(t)$ prevailing at t has a certain payoff associated to it, as reflected by the cooperation payoff $\zeta_{ij}(t)$ in the stage PD game played by i and j. To evaluate the stability of a network $g(t)$, one needs information on the overall attainable payoffs. This suggests identifying the state of the system prevailing at any given t with the pair $\omega(t) = \{[\vec{g}_i(t)]_{i \in N}, [\zeta_{ij}(t)]_{ij \in g(t)}\}$, the only restriction being that each $\vec{g}_i(t)$ should consist of at most m elements and every $\zeta_{ij}(t)$ belongs to the suitable range of payoffs (see below).

The social dynamics defining the law of motion for the current state embodies three distinct components: payoff update, link formation plus search, and removal of unstable links. Assuming, for simplicity, that they operate in that order, I separately describe each of these components.

2.2.1. Payoff update

First, suppose that the payoff of each link may be subject to a random update of its associated payoff. More precisely, with some independent probability ε, every link $ij \in g(t)$ has its payoff changed from $\zeta_{ij}(t)$ to $\zeta'_{ij}(t)$, where the latter is drawn afresh from some non-negative real interval $[\underline{\zeta}, \overline{\zeta}]$ according to a stationary (and common) probability distribution with continuous density f_ζ. For future reference, denote by $\omega'(t) = \{[\vec{g}_i(t)]_{i\in N}, [\zeta'_{ij}(t)]_{ij\in g(t)}\}$, the new state thus generated.

2.2.2. Link formation and search

Next, let us suppose that players may enjoy two alternative (say, exclusive) routes of search and consequent formation of fresh links: component-bound 'local' search and unrestricted 'global' search. Whereas the first route may be conceived as the more common way of accessing new information (since it is mediated by the social network), the second one could be regarded as more extraordinary (and thus only occasional). Formally, let us posit that every player $i \in N$ enjoys a local search option with independent probability p, whereas she receives a global one with probability $(1-p)q$. Then, if one of these alternative options does materialize, the corresponding events that take place are as follows.

2.2.2.1. Local search. Given $g(t)$, let $M_i(g(t))$ represent the set of players who are *not* direct neighbors of i but belong to the same *component* as i – i.e. there is a path in $g(t)$ joining them to node i.[6] These are the potential new partners of player i when she receives a local (component-bound) search opportunity. In this event, what happens is that she observes fresh and independent payoff draws ζ_{ij} (according to the probability density f_ζ) for all $j \in M_i(g(t))$. On the basis of this information, player i is allowed to establish *one* new link, possibly removing one of her pre-existing links if she already supports the maximum number of m links.

More specifically, what I postulate is that player i chooses a new link with some $j \in M_i(g(t))$ if, and only if, both of the following conditions hold:

(i) The payoff drawn afresh $\zeta_{ij} \geqslant \zeta_{ik}$ for all $k \in M_i(g(t))$.
(ii) If $|N_i(t)| = m$, $\zeta_{ij} > \zeta_{i\ell}$ for some $\ell \in N_i(t)$.

Thus, the creation of new links is supposed *gradual*, i.e. at most one at a time is formed. In case that (i) applies for more that one $j \in M_i(g(t))$, any of them is selected with the same probability. On the other hand, if (ii) applies, then the link $i\ell$ to the current neighbor ℓ with the lowest payoff is removed. Again, if several of those exist, each of them is chosen with equal probability.

2.2.2.2. Global search. If some player i receives a revision opportunity through global search, she gets the possibility of forming one new link with another randomly

[6] Alternatively, one could postulate (as in Marsili et al. (2004)) that only those agents in the component who are at a certain distance from player i are possible candidates for new partners. I have explored variations of the model parametrized by different search radii and the essential conclusions are unaffected. Thus, for simplicity, I abstract from these considerations here.

selected individual j, possibly not in her component, all individuals in the population selected with equal probability. Again, associated to this new potential partner a payoff ζ_{ij} is drawn according to the probability density f_ζ. Then, as above, the link is established, subject to the possible removal of a pre-existing link if the number of those links is maximum and one of them has a lower payoff.

2.2.3. Removal of pairwise-unstable links

Let $\omega''(t) = \{[\vec{g}''_i(t)]_{i\in N}, [\zeta''_{ij}(t)]_{ij\in g(t)}\}$ be the state induced by the two former components of the law of motion. Then, for every link $ij \in g''$, let the players i and j involved in the link evaluate whether both of the following incentive-compatibility conditions hold:

$$\pi''_i(g'' \overset{ij}{\setminus} C) \geqslant \pi''_i(g'' \overset{ij}{\setminus} D),$$
$$\pi''_j(g'' \overset{ij}{\setminus} C) \geqslant \pi''_j(g'' \overset{ij}{\setminus} D).$$

If either of these conditions is violated, the link ij is judged unstable by the players involved and thus is removed. Once such a check of pairwise-stability has been completed for all links in g'', let $\omega'''(t)$ refer to the resulting state where only the links that have been assessed as pairwise-stable remain. This state is then carried over to the next period, by making $\omega(t+1) = \omega'''(t)$.

3. Static analysis: characterization of pairwise-stability

Organizing the analysis along the same steps followed to present the model, it is useful to start with a static (equilibrium) approach, then turning to a full dynamic analysis of the network formation process. The first task, therefore, is to understand the implications of pairwise stability on the maintenance of links. This is carried out by Propositions 1 and 2 below, which provide an intuitive analogous characterization of PSN. This characterization hinges upon a certain measure of cohesiveness of the network, that reflects the average (geodesic) distance separating the neighbors of the different players.

To introduce such measure of network cohesiveness formally, let us define, for any given player $i \in N$ and any two of her neighbors $j, k \in N_i$, the *i-excluding distance* between j and k, denoted by $d^i(j,k)$, as the length of the shortest path joining j and k which does *not* involve player i. The interpretation of this distance is straightforward: it is the number of steps (and therefore periods, in the repeated game) which would be required for any information held by j (or k) to reach k (or j) without the concourse of player i. As usual, it is postulated that $d^i(j,j) = 0$ for any $j \in N_i$, while if no i-excluding path exists between k and j it will be convenient to posit that $d^i(j,k) = \infty$.

First, I state the result characterizing pairwise-stability of a network in the stronger sense:

Proposition 1. *Consider any network $g \subset N \times N$ and let $[\zeta_{ij}]_{ij\in g}$ stand for the cooperation payoffs that can be earned for each of its links. Then, g is a PSN if, and* *only if, for all* $ij \in g$:

$$\zeta_{ij} \geqslant \frac{1-\delta}{\delta}\, v - \sum_{k \in N_i \setminus \{j\}} \zeta_{ik} \delta^{d^i(j,k)}. \tag{2}$$

The proof of this result is provided in the appendix. To understand its implications note that, in the absence of network effects, the condition required for the stability of a network g would be that, for every link $ij \in g$,

$$\zeta_{ij} \geqslant \frac{1-\delta}{\delta}\, v. \tag{3}$$

Thus, in contrast with the above condition, we observe that the network-based strategic effects embodied by (4) may enlarge substantially the circumstances under which cooperation (and thus stability of the network) can be attained. Turning now to the stronger version of pairwise-stability, we have the following counterpart of the former result, whose proof is also relegated to the appendix.

Proposition 2. *Consider any network* $g \subset N \times N$ *and let* $[\zeta_{ij}]_{ij \in g}$ *stand for the cooperation payoffs that can be earned for each of its links. Then,* g *is a Weakly Pairwise-Stable Network (WPSN) if, and only if, for all* $ij \in g$:

$$\zeta_{ij} \geqslant \frac{1-\delta}{\delta}\, v - \sum_{k \in N_i \setminus \{j\}} \left(\zeta_{ik} - \frac{1-\delta}{\delta}\, v \right) \delta^{d^i(j,k)}. \tag{4}$$

Note that, as expected, (4) implies (2), a mere reflection of the fact that the wider range of deviations allowed under the SPSN concept leads to a stronger version (i.e. a refinement) of the original notion of pairwise stability. The two conditions, however, display a similar form and also lead to analogous insights. Both indicate that any two players, i and j, will have their incentives for cooperation enhanced the higher number and more valuable neighbors each has who are not too far from the potential partner.

4. Network dynamics

4.1. Ergodicity and the role of volatility

First, as a benchmark case, it is useful to consider the extreme scenario where *no* fresh links are ever formed through search (i.e. $p = q = 0$) and prevailing payoffs are *never* subject to update ($\varepsilon = 0$). In this context, the dynamics reduces to a mere chain of link removals, as the links which are deemed unstable are being withdrawn by the agents involved. Eventually, the process reaches a stationary state, where the induced network is pairwise stable in the sense of Definition 1. In fact, it is obvious that the sets of pairwise-stable and stationary networks coincide in this case, which thus provides a simple dynamic framework to understand the PSN concept.

But, of course, our primary interest pertains to the richer context where $p, q, \varepsilon > 0$. A first basic step in the analysis of this full-fledged scenario is the establishment of conditions under which the induced Markov process may be guaranteed to be ergodic. As explained, this property provides the theoretical basis for the use of numerical simulations in unveiling the long-run (dynamically invariant) behavior of the process.

If we made $\varepsilon > 0$ but had $q = p = 0$, then volatility would face no countervailing force and the process would eventually become absorbed by the empty network. This would render the process ergodic but trivially so. In fact, the same long-run state of affairs would prevail if only $q = 0$ but $p > 0$. For, in this case, volatility would again succeed in eventually removing *all* links and local search would be incapable of allowing the system to escape this situation. This motivates postulating that $q > 0$. In fact, it is not difficult to see that global search alone would be enough to obtain non-trivial ergodicity in the presence of volatility. We shall be interested, however, in studying scenarios where $p > 0$ as well since, in the present context, it is natural to have the current network bear significantly on the search for new links – indeed, our simulations will focus on cases where p is much larger than q, thus attributing to local search the preeminent role in network formation.

To address matters formally, recall that $f_\zeta(\cdot)$ stands for the probability density which governs every fresh draw of payoffs, whose support is given by a non-negative interval $[\underline{\zeta}, \overline{\zeta}]$. Further recall that $(1 - \delta)/\delta v$, is the payoff threshold that marks the bilateral supportability of cooperation between a pair of 'isolated' agents (cf. (3)). Bearing this in mind, the following result simply states that ergodicity is guaranteed as long as, for any cooperation payoff ζ_{ij} redrawn afresh, there is some prior uncertainty as to whether it may be supported on a strictly bilateral fashion.

Proposition 3. *Assume* $\underline{\zeta} < (1 - \delta)/\delta v < \overline{\zeta}$ *and* $\varepsilon, p, q \in (0, 1)$. *Then, provided the population N is large enough, the social dynamics described in Subsection 2.2 is governed by an ergodic stochastic process.*

Proof. See appendix. □

The former result requires both the concourse of search *and* volatility in order to attain the independence of initial conditions that is the mark of ergodicity. In this respect, it is instructive to consider what would be the role of search if, unlike what is posited above, there is no volatility to struggle with. A clear and intuitive picture of the induced situation is provided by the following result.

Proposition 4. *Assume* $\underline{\zeta} < (1 - \delta)/\delta v < \overline{\zeta}$ *and* $p, q \in (0, 1)$. *Then, if* $\varepsilon = 0$, *the social dynamics leads almost surely to a path where the network reaches the maximum connectivity and the induced total payoff (aggregated over the whole population) converges to it maximum value* $2mn\overline{\zeta}$.

Proof. See appendix. □

The previous result indicates that, in the absence of payoff volatility, the accumulation of social capital must eventually reach a maximum level. This serves to highlight the key role to be played by ε as the leading parameter of the ensuing discussion.

4.2. Numerical analysis

Building upon the results presented in the previous subsection, we now conduct an exhaustive analysis of the model based on numerical simulations. Our main objective is to understand how the interplay between environmental volatility and the dynamics of the social network shapes the long-run performance of the system. For the sake of focus, the simulation scenario to be considered here displays rather stringent payoff conditions, so that the essential points may arise in a starker fashion. However, a wide range of parameter configurations have also been considered, yielding analogous conclusions.

The simulation setup involves $100(=n)$ individuals who display a common discount rate $\delta = \frac{3}{4}$ and interact according to an IRPD game with *stage* payoffs as given by the payoff table in (1) for $v = 4$. The cooperation payoffs ζ_{ij} are drawn, randomly and independently, according to a uniform distribution over the interval $[\underline{\zeta}, \overline{\zeta}] = [0.4, 1.4]$. Thus, as required by Proposition 3, the threshold $[v(1-\delta)]/\delta = 4/3$ which marks the possibility of supporting cooperation bilaterally (cf. (3)) belongs to the payoff support and thus ergodicity is guaranteed. Payoff conditions, however, are rather 'tight' in that the probability $\Pr[\zeta_{ij} \geqslant [(1-\delta)/\delta]v]$ for a fresh draw to exceed the aforementioned threshold is just $\frac{1}{15}$.

Finally, we postulate that the (independent) probability with which an individual receives a revision opportunity based on local search is equal to $p = 0.1$, while in case no such opportunity arises the conditional probability for enjoying an instance of global search is $q = 0.01$. The maximum number of links that any given player can actively support is set equal to $m = 2$ (therefore, the average network degree is at most 4).

In the above described scenario, the numerical analysis is divided in two parts. First, in Subsection 4.2.1, I consider the leading case where players may take full advantage of network effects in supporting cooperation – that is, strategic behavior is 'network-based' and pairwise stability is given by Definition 1. Then, in Subsection 4.2.2, I turn to studying how matters are affected if players do *not* rely on such network effects and all cooperation must be supported bilaterally.

4.2.1. Network-based strategic behavior

The analysis focuses throughout on the four variables contemplated in our earlier discussion: *network density* (average node degree), *network cohesiveness* (average neighbor distance), *network span* (relative size of the largest components), and *payoff performance* (average payoff earned per link).

By virtue of ergodicity, we know that, a.s. along *any* sample path (i.e. independently of initial conditions), the empirical mean of each of the above variables converges to the expected value induced by the unique invariant measure of the process. To be more precise, denote by μ the unique invariant of the ergodic process and let $x(\omega)$ be the value for anyone of those variables when the state is ω. Then, ergodicity implies that, with probability one over the possible sample paths $\{\omega(t)\}_{t=0}^{\infty}$, we have

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} x(\omega(t)) = \int x(\omega)\,\mathrm{d}\mu,$$

Indeed, this convergence is confirmed by our numerical simulations, which show that the empirical means of the aforementioned variables converge, independently of the initial conditions, to common long-run values. Our main objective is to understand how these long-run invariant values depend on the leading parameter of our analysis, the volatility rate ε. The results are summarized in Figs. 1(a)–(d).

First, Fig. 1(a) shows that payoff volatility has a negative effect on the density of the social network. Intuitively, this follows from the fact that, as ε rises, there is a larger fraction of existing links that lose its pairwise stability due to payoff update. This phenomenon may be heuristically understood as a sort of depreciation of the accumulated social capital. It reflects, in effect, a negative drift on the stock of existing links that have been gradually accumulated over time through successful search.

Next, Fig. 1(b) depicts what is possibly one of the most remarkable regularities found in the analysis. It shows that, as ε grows, the social network endogenously adapts to increased volatility by becoming more *cohesive* in the long run – i.e. the average neighbor distance falls.[7] This is in line with what has been learned so far about the role played by network effects in enhancing the incentives for cooperation (recall, for example, the characterization of pairwise-stability in Proposition 1). Thus, in this sense, what we find is that the social dynamics responds to increased volatility by *endogenously* building up the strength of these network effects. Agents, unwittingly, adjust myopically their links over time so that the social network ends up adapting to higher volatility by raising its long-run cohesiveness. This, in the end, has the beneficial (but unintended) effect of deterring opportunistic behavior more effectively as such a network-based deterrence becomes more critical.

Finally, it is worth emphasizing that the aforementioned considerations are to be regarded as quite strong. For, in particular, they are strong enough to prevail over the general decrease in connectivity that results as ε grows (recall Fig. 1(a)). By itself, such a decrease in overall connectivity tends to increase path lengths and thus would entail, absent the considerations explained, a consequent increase in average neighbor distance.

Fig. 1(c) describes the long-run effect of payoff volatility on the size of the largest two components. The first interesting point to observe is that, independently of ε, most of the connected players belong to a *single* major component, the second-largest component remaining very small throughout. This conclusion is somewhat

[7]For certain nodes, the distances between some of their neighbors could be infinite, i.e. there could be no path joining them other than the one passing through the node in question. To avoid this problem in computing average distances, I truncate the neighbor distance in those cases to be equal to n, just above the maximum possible *finite* distance of $n-1$. This convention is of minor actual relevance and does not affect the main gist of the analysis.

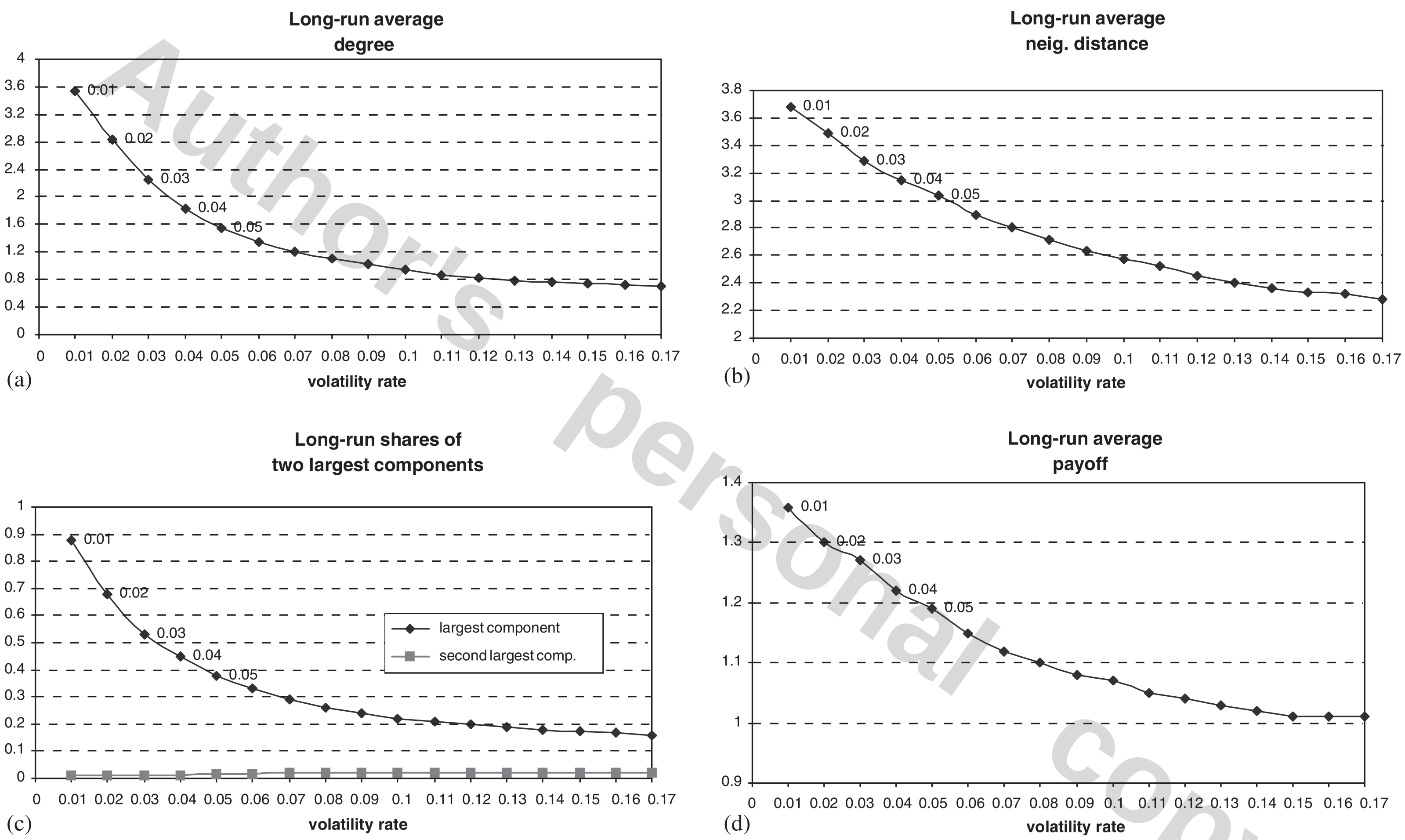

Fig. 1. The effect of volatility on the long-run values of connectivity (average degree), cohesiveness (average neighbor distance), relative size of the two largest components, and the average payoff *per* link.

reminiscent of the well known results of Theory of Random Graphs which assert that, beyond a certain 'connectivity threshold,' there arises a *single* giant component in the graph. (These mathematical results, however, do not appear to readily applicable here since, in our case, the formation of the social network is very much the result of a non-random mechanism.) On the other hand, concerning the effect of ε on the relative size of the largest component, the expected negative dependence is observed, which is another reflection of the detrimental effect of volatility on the stock of social capital that was already discussed in connection to Fig. 1(a).

As a further graphical manifestation of this phenomenon, Fig. 2 also depicts the networks prevailing at the end of the time horizon for two contrasting scenarios: one where payoff volatility is low ($\varepsilon = 0.02$) and another where it is relatively high ($\varepsilon = 0.12$). These networks provide a tangible confirmation that the long-run network structure is indeed polarized (i.e. a single giant component coexists with a more or less sizable set of disconnected players). They also illustrate that, as explained above pertaining to the cohesiveness of the network, higher volatility leads to lower neighbor distances – this is heuristically suggested by the existence of a significant number of paths with 'loose' end nodes for the lower $\varepsilon = 0.02$, while no such paths exist for $\varepsilon = 0.12$.[8]

Finally, let us turn to Fig. 1(d), which describes the negative effect of payoff volatility on the long-run average payoff earned per interaction (or link). The evidence gathered here serves as an interesting complement to that displayed in Fig. 1(a). It shows that, as ε grows, there is not only a decrease in the network connectivity (i.e. games being played) but also in the average payoff earned *per link*. The decline in network density was formerly explained as a kind of social-capital depreciation that renders a growing fraction of preexisting links unstable. In contrast, the present negative effect on the average payoff earned per existing link should be largely understood as a consequence of the narrower scope of search brought about by the decreasing size of the network's largest component. This more confined (and thus less effective) search deteriorates, even when successful, the expected level of attainable payoffs.

4.2.2. Network-free strategic behavior

To shed further light on the results obtained for our leading scenario in the previous section, it is most instructive to compare it with one devoid of network-based strategic effects. Thus suppose that players' strategic behavior no longer is subject to these effects and therefore the condition for cooperation between any two pair of connected players i and j is (3). This may be understood as reflecting a different norm or convention used in the society – one where players react to each of their partners only according to the information gathered on their corresponding bilateral play. Then, strategically speaking, every pairwise interaction is to be regarded as fully independent of any other. This leaves the social network with the

[8]To be precise, it turns out that the neighbor distance of the first network ($\varepsilon = 002$) is 3.6 whereas that of the second ($\varepsilon = 0.12$) is 2.7.

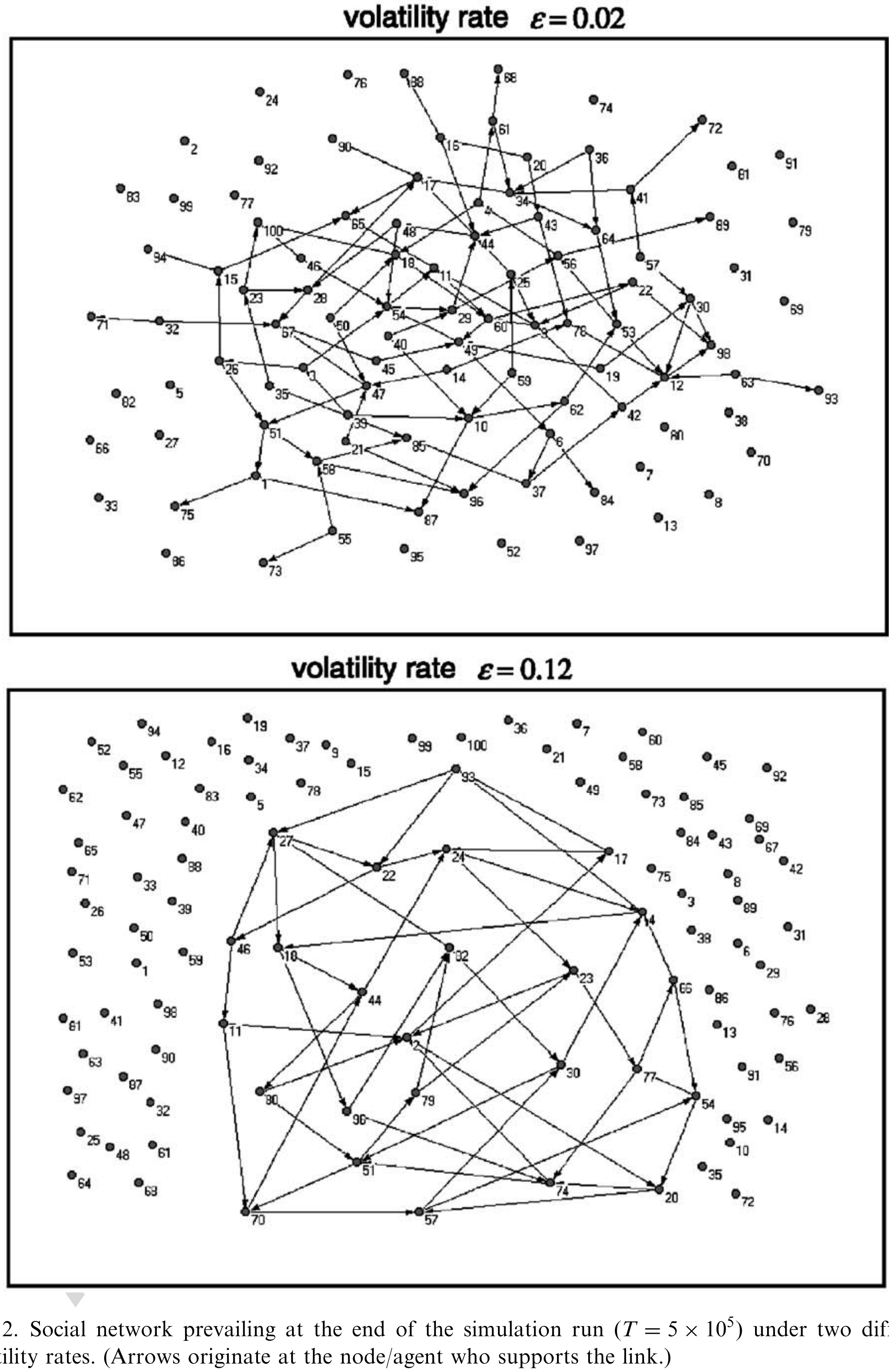

Fig. 2. Social network prevailing at the end of the simulation run ($T = 5 \times 10^5$) under two different volatility rates. (Arrows originate at the node/agent who supports the link.)

only, but yet important, role of defining the channels through which the information diffuses and the population conducts ordinary ('local') search.

Rather than pursuing a completely parallel analysis to that undertaken before, let us restrict attention to only two of the variables where the contrast is more acute and interesting: average node degree (network density), and average neighbor distance (network cohesiveness). By relying on ideas analogous to those used in the proof of Proposition 3, it is straightforward to show that the induced stochastic process is also ergodic in this case and thus the long-run values for these variables are uniquely defined. Under the same underlying parameters as above, their dependence on ε is displayed in Fig. 3(a)–(b).

The main points of contrast with the leading network-based scenario can be summarized as follows.

1. The bite on long-run network density caused by increased volatility manifests itself at much lower rates than in the original scenario. Specifically, the average degree starts to face a significant fall at values of ε that are *one order of magnitude* smaller than before.

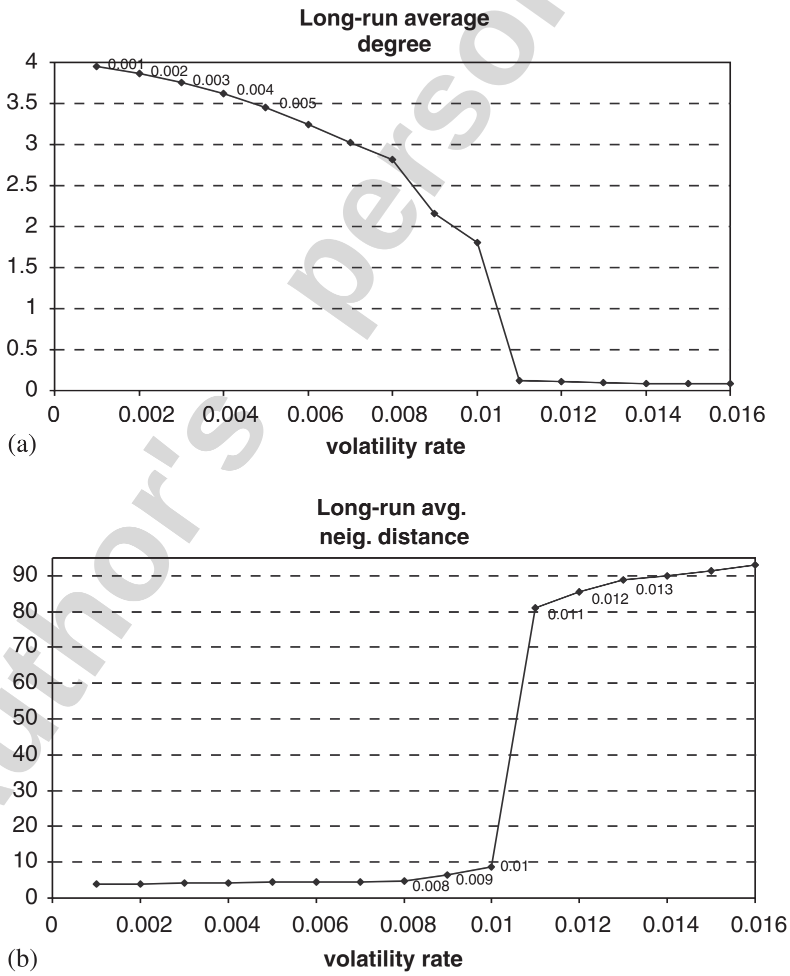

Fig. 3. The effect of volatility on the long-run values of connectivity (average degree) and cohesiveness (average neighbor distance) in a scenario without network-based strategic effects.

2. An additional, and very telling, difference concerns the abruptness now observed in the transition from a regime with high network density to a regime with a sparse social network. In the absence of network strategic effects, this transition takes place quite sharply, in at least two complementary senses. On the one hand, the full change is essentially completed within a range for ε that, again, is shorter than in the original framework by an order of magnitude. On the other hand, the shape of the 'curve' which traces this change is essentially concave (i.e. displays increasing differences) along the transition. This contrasts with the behavior displayed under network strategic effects, where changes in volatility lead to a gradual effect on long-run density.
3. The behavior of cohesiveness in the present case (as measured by the average neighbor distance) also displays a marked contrast with the results obtained in the leading scenario with network strategic effects. There, we stressed that an increase in ε leads to rising network cohesiveness, which is how the social dynamics endogenously adapts to, and partially offsets, the entailed negative effects. Now, however, cohesiveness entails no relevant payoff consequences, and therefore an increase in ε induces a corresponding rise in neighbor distance. As explained in Subsection 4.2.1, this may be simply interpreted as a reflection of the fact that, absent other considerations, lower connectivity tends to increase any measure of graph distance – also average neighbor distance.

5. Mean-field analysis

In this section, the objective is to complement the numerical analysis reported above with a simplified mean-field study of the model. Mean-field methods are commonly used by statistical physics in the study of large complex systems of interacting entities.[9] In a nutshell, what underlies the approach is the presumption that, in the presence of a large number of stochastic interactions, the aggregate behavior of the system can be reasonably well understood (i.e. 'modelled') in expected terms. This, in essence, amounts to positing a *deterministic* representation of the process that disregards fluctuations and only depends on the average (or expected) magnitudes of the variables involved.

The mean-field description of the model proposed here is particularly simple, since our limited aim is to shed light on the sharp contrast between the network-based and network-free scenarios considered in Subsections 4.2.1 and 4.2.2. We consider, therefore, a mean-field formulation of the dynamics that focuses on the evolution of network connectivity, while other features of the process (for example, the effect of network architecture on search) are accounted for in a reduced-form fashion. As indicated, for an illustration of how a richer mean-field analysis may encompass other network features (e.g. network cohesiveness or the full degree distribution), the reader is referred to Marsili et al. (2004) where this approach is applied in detail to a non-strategic scenario.

[9]See, for example, the classical monograph by Stanley (1971).

To facilitate the exposition, it is useful to decompose the law of motion for the *mean-field dynamics* (MFD) in three stages. First, I formalize the mechanism of link removal induced by the requirement of pairwise stability. Second, I specify how this mechanism operates in the presence of payoff volatility. Lastly, I introduce the process of search and creation of new links.

(a) *Link removal*: Given any possible link $ij \in N \times N$, let $\lambda_{ij}(t) \in \{0, 1\}$ specify whether this link is in place at t (an event which is signified by $\lambda_{ij}(t) = 1$) or not ($\lambda_{ij}(t) = 0$). Then, for this link ij, the dynamics of link-removal (due to pairwise instability) may be formulated as follows:

$$\lambda_{ij}(t+1) = \lambda_{ij}(t) \cdot \mathscr{I}[\zeta_{ij}(t+1) - \beta_{ij}(t+1)], \tag{5}$$

where

- $\zeta_{ij}(t+1)$ is the cooperation payoff prevailing at $t+1$ for the game played by i and j (it may normalized to zero when no link connects i and j).
- $\mathscr{I}[\cdot]$ is an indicator function defined by $\mathscr{I}[y] = 0$ if $y < 0$ and $\mathscr{I}[y] = 1$ if $y \geqslant 0$.
- $\beta_{ij}(t+1)$ is the threshold for the cooperation payoff ζ_{ij} that is required – given the prevailing network and other payoffs (cf. (2)) – for the pairwise stability of the link ij, i.e. $\beta_{ij}(t+1) = \max\{\varphi_{ij}(t+1), \varphi_{ji}(t+1)\}$, where

$$\varphi_{ij}(t+1) = \frac{1-\delta}{\delta} v - \sum_{k \in N_i(t) \setminus \{j\}} \delta^{d^i(k,j)(t+1)} \left[\zeta_{ik}(t+1) + \frac{1-\delta}{\delta} v \right] \tag{6}$$

and $\varphi_{ji}(t+1)$ is defined reciprocally.

(b) *Payoff volatility*: Combining (5) with the process of payoff update that affects each link ij with probability ε, the expected value of its link indicator satisfies:

$$\mathbf{E}_t[\lambda_{ij}(t+1)] = (1-\varepsilon)\lambda_{ij}(t) + \varepsilon\lambda_{ij}(t) \int_{\beta_{ij}(t)}^{\bar{\zeta}} f_\zeta(z)\,\mathrm{d}z, \tag{7}$$

where f_ζ is the continuous density that governs every payoff draw and the assumption is made that the (point) expectation at t on $\beta_{ij}(t)$ is 'static,' i.e. $\beta_{ij}(t+1) = \beta_{ij}(t)$. Under those conditions, (7) simply embodies the idea that any existing link ij which is subject to payoff update is expected to remain in place with probability $\Pr\{\zeta_{ij} \geqslant \beta_{ij}(t)\}$, while those whose payoff is not updated are expected to stay.

(c) *Search and link creation*: As explained, this component of the dynamics is introduced in 'reduced form.' Thus I posit that the expected increase in the number of links per node is given by some differentiable function $\phi(\kappa)$ of the current average degree $\kappa \in [0, 2m]$. (Recall that m is the maximum number of links supportable by any given individual and therefore $2m$ is the maximum value of κ.) Given the features displayed by local search (recall Subsection 2.2), the function $\phi(\cdot)$ must verify the following boundary conditions:

$$\phi(0) = \eta \int_{\frac{1-\delta}{\delta}v}^{\bar{\zeta}} f_{\zeta}(z)\,\mathrm{d}z, \tag{8}$$

$$\phi(2m) = 0. \tag{9}$$

The first condition indicates that, when there are no links available, only global search can contribute (if successful) to the creation of new links. On the other hand, the second condition simply follows from the fact that no more links can be created when agents are already at their maximum supporting capacity. Finally, in addition to (8)–(9), it is natural to postulate that, below the maximum capacity $2m$, search yields some expected increase in connectivity, i.e.

$$\phi(\kappa) > 0, \quad \forall \kappa < 2m$$

while, at low connectivity, we have

$$\phi'(0) = 0, \tag{10}$$

i.e. vanishing marginal effects when there are almost no links. The latter condition has been explicitly derived from a detailed formalization of local search in Marsili et al. (2004). It is intuitive since when only very few links are in place, the network is composed of isolated pairs and thus local search is fruitless.

Combining (a)–(c) above, the key approximation step undertaken by the mean-field approach is twofold. Firstly, it identifies expected and actual motion, so that the average degree at $t+1$, denoted by $\kappa(t+1)$, is made equal to

$$\kappa(t+1) \equiv \mathbf{E}_t \left[\frac{1}{n} \sum_{i \in N} \sum_{j \neq i} \lambda_{ij}(t+1) \right].$$

Secondly, it replaces the link-specific magnitudes in (7) – in particular, each $\beta_{ij}(t)$ – by an average $\hat{\beta}(t)$ computed across all links ij prevailing at t. This leads to the following difference equation governing the average degree:

$$\begin{aligned} \kappa(t+1) &= \left[(1-\varepsilon) + \varepsilon \int_{\hat{\beta}(t)}^{\bar{\zeta}} f_{\zeta}(z)\,\mathrm{d}z \right] \left[\frac{1}{n} \sum_{i \in N} \sum_{j \neq i} \lambda_{ij}(t) \right] + \phi(\kappa(t)) \\ &= \left[1 - \varepsilon + \varepsilon \left(1 - \int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z)\,\mathrm{d}z \right) \right] \kappa(t) + \phi(\kappa(t)) \\ &= \left[1 - \varepsilon \int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z)\,\mathrm{d}z \right] \kappa(t) + \phi(\kappa(t)) \end{aligned}$$

which, for analytical tractability, is convenient to transform into the analogue differential equation:

$$\dot{\kappa}(t) = -\varepsilon \left[\int_{\underline{\zeta}}^{\hat{\beta}(t)} f_{\zeta}(z)\,\mathrm{d}z \right] \kappa(t) + \phi(\kappa(t)), \tag{11}$$

where the parameters ε and η, or the value of the function $\phi(\cdot)$ are now interpreted as rates in continuous time.

The above general formulation is applicable whether network-based strategic effects are present or not. The only difference between these two alternative scenarios resides in how the threshold $\hat{\beta}(t)$ is specified. Thus, if network effects are absent, this feature is simply captured by making $\hat{\beta}(t)$ *constantly* equal to $v[(1-\delta)/\delta]$ (cf. (3)). The dynamics (11) then is particularized as follows:

$$\dot{\kappa}(t) = -\varepsilon\left[\int_{\underline{\zeta}}^{v\frac{(1-\delta)}{\delta}} f_{\zeta}(z)\,\mathrm{d}z\right]\kappa(t) + \phi(\kappa(t)). \tag{12}$$

To see how this dynamics reflects the sharp phase transition observed in the simulations, suppose that, as we have maintained throughout, $\underline{\zeta} < v[(1-\delta)/\delta] < \overline{\zeta}$. Then, first note that for sufficiently small ε, (12) yields a unique globally stable state arbitrarily close to $\kappa_{\max} = 2m$. On the other hand, I now argue that as ε rises, there is a threshold on ε which marks a discontinuity in the behavior of the dynamics akin to that observed in the simulations. To make this point in the starkest manner, it is convenient to focus on the extreme case where $\eta = 0$ and thus $\phi(0) = 0$. (The reader can easily check that an analogous conclusion would be reached for small positive η.)[10] Denote by $\hat{\kappa}(\varepsilon)$ the highest equilibrium degree which is asymptotically (locally) stable according to the mean-field dynamics (12). Then, we shall establish the following claim.

Claim 1. *Consider the MFD* (12) *representing the process under no network-based strategic effects with* $\eta = 0$. *Then, there are some* $\bar{\varepsilon}$ *and* $\bar{\kappa}$, *both strictly positive, such that*

$$\varepsilon < \bar{\varepsilon} \;\Rightarrow\; \hat{\kappa}(\varepsilon) \geqslant \bar{\kappa}, \tag{13}$$

$$\varepsilon \geqslant \bar{\varepsilon} \;\Rightarrow\; \hat{\kappa}(\varepsilon) = 0. \tag{14}$$

Proof. To verify this claim, first note that a certain state κ^* is asymptotically stable for (12) if, and only if,

$$\varepsilon\left[\int_{\underline{\zeta}}^{v\frac{(1-\delta)}{\delta}} f_{\zeta}(z)\,\mathrm{d}z\right]\kappa^* > \phi'(\kappa^*). \tag{15}$$

Then observe that a necessary and sufficient condition for (15) to hold at some interior $\kappa^* > 0$ is (cf. Fig. 4):

$$\varepsilon\left[\int_{\underline{\zeta}}^{v\frac{(1-\delta)}{\delta}} f_{\zeta}(z)\,\mathrm{d}z\right] < \max_{\kappa\in[0,2m]} \frac{\phi(\kappa)}{\kappa}.$$

[10] In that case, the statement that $\hat{\kappa}(\varepsilon)$ falls discontinuously to zero as ε exceeds the threshold would be replaced by the assertion that $\hat{\kappa}(\varepsilon)$ is an infinitesimal in η (i.e. converges to zero as $\eta \downarrow 0$). Here, of course, we implicitly assume that $\phi(\kappa)$ is jointly continuous in both κ and the underlying η.

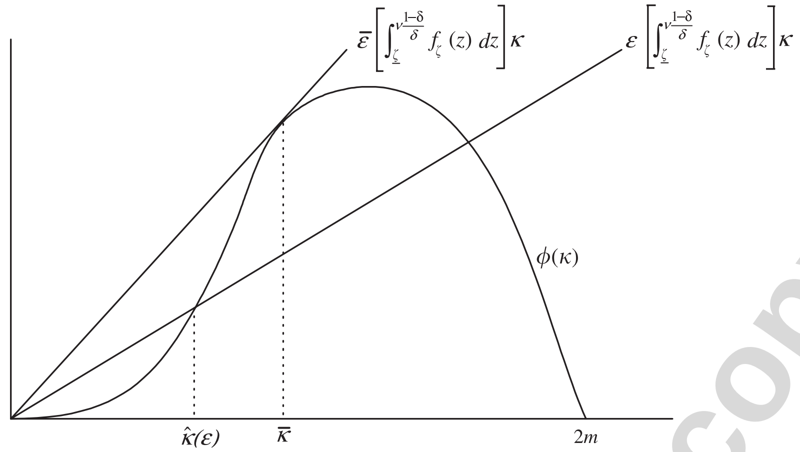

Fig. 4. Equilibria for the mean field dynamics in the absence of network-based strategic effects with $\bar{\varepsilon}$ given by (16) and $\varepsilon > \bar{\varepsilon}$.

Hence it follows that the threshold $\bar{\varepsilon}$ contemplated in (13)–(14) is given by

$$\bar{\varepsilon} = \left[\int_{\underline{\zeta}}^{v\frac{(1-\delta)}{\delta}} f_{\zeta}(z)\,\mathrm{d}z\right]^{-1} \max_{\kappa\in[0,2m]} \frac{\phi(\kappa)}{\kappa} \tag{16}$$

which is always well-defined, in view of (10). In turn, $\bar{\kappa}$ may be readily defined as follows:

$$\bar{\kappa} \equiv \min_{\kappa} \arg\max_{\kappa} \frac{\phi(\kappa)}{\kappa},$$

that is certain to be positive, again by virtue of (10). □

Conditions (13)–(14) imply that whether or not the MFD can sustain in a robust (i.e. locally stable) fashion some *positive* network density depends *discontinuously* on the volatility parameter. This is in line with the numerical simulations reported in Subsection 4.2.2 if we conceive the mean-field approach as reflecting the 'dominant forces' at work in the original stochastic process. In this perspective, the (deterministic) MFD is to be viewed as *singling out* what behavior of the original (stochastic) process may be observed with some persistence over time in the presence of stochastic fluctuations. That is, only the counterparts of those states that are locally stable in the MFD may have any significance (i.e. weight) in the long-run dynamics of full-fledged stochastic process. This then suggests using (13)–(14) to 'predict' that as the volatility rate ε exceeds some threshold, the long-run behavior of the process should witness an abrupt decrease in network density to very low levels – i.e. the behavior indeed observed in Fig. 3(a).

By way of contrast, let us now turn our attention to the polar scenario where network effects do impinge on agents' strategic behavior. The implication of this feature on the MFD formulation given by (11) is that $\hat{\beta}(t)$ is no longer constant but has to be assumed dependent on $\kappa(t)$ – the state of the system – through a certain function $h(\cdot)$. This leads to a particularization of (11) as follows:

$$\dot{\kappa}(t) = -\varepsilon \left[\int_{\underline{\zeta}}^{h(\kappa(t))} f_{\zeta}(z)\, \mathrm{d}z \right] \kappa(t) + \phi(\kappa(t)). \tag{17}$$

In view of (6) and (7), it is natural to postulate that this function should be decreasing and satisfy:

$$h(0) = \frac{1-\delta}{\delta}\, v \tag{18}$$

and[11]

$$h(2m) < \underline{\zeta}. \tag{19}$$

We now show that, in contrast with Claim 1 above, the induced mean-field dynamics is in essential accord with the behavior observed in Subsection 4.2.1 under network strategic effects.

Claim 2. *Consider the MFD* (17) *representing the process under no network-based strategic effects with* $\eta = 0$. *Then, both* $\kappa = 0$ *and* $\kappa = 2m$ *are locally stable points.*

Proof. Denote by $F(\kappa) \equiv -\varepsilon[\int_{\underline{\zeta}}^{h(\kappa)} f_{\zeta}(z)\, \mathrm{d}z]\kappa + \phi(\kappa)$ the function governing the dynamics in (17). Then, (18) and (19), in combination with (10), readily imply that there are some (relative) neighborhoods of $\kappa = 0$ and $\kappa = 2m$, W_0 and W_{2m} respectively, such that

$$\kappa \in W_0 \backslash \{0\} \;\Rightarrow\; F(\kappa) < 0, \tag{20}$$

$$\kappa \in W_{2m} \backslash \{2m\} \;\Rightarrow\; F(\kappa) > 0 \tag{21}$$

for all $\varepsilon > 0$. This establishes the desired conclusion. □

Thus, if we rely again on the notation $\hat{\kappa}(\varepsilon)$ to denote the highest equilibrium degree which is asymptotically stable according to the corresponding MFD, Claim 2 implies that $\hat{\kappa}(\varepsilon) = 2m$ for all $\varepsilon \geqslant 0$. This suggests that an abrupt fall in connectivity to very low levels does not happen in the original process when players's behavior is affected by network effects – indeed the conclusion obtained in the numerical simulations.[12]

What explains the contrast between the network-based and the network-free scenarios? The mean-field description of the process provides a quite transparent

[11]Condition (19) implicitly presumes that m is large enough and that, as it seems intuitive, average payoffs do not decrease, or neighbor distances increase, with connectivity. In fact, this condition is satisfied in the simulations reported in Subsection 4.2 when the average network degree is high. Then, the average φ_{ij} given by (6) – and thus the average β_{ij} – is significantly lower than $\underline{\zeta}$.

[12]In fact, one can further specialize the MFD with additional assumptions (e.g. linearity of $h(\cdot)$ or uniformity of the payoff distribution) to conclude that $\kappa = 0, 2m$ are the only asymptotically stable states under network effects – cf. Vega-Redondo (2002). In this case, there is of course an additional unstable equilibrium $\check{\kappa}(\varepsilon) \in (0, 2m)$ which separates the basins of attraction of the other two, and which happens to vary continuously with ε.

answer. When network effects are present, an agent with high connectivity also tends to have, on average, a local interconnected network architecture which deters opportunistic behavior, even when a few of her links are subject to a negative redraw of its payoff. This is, in essence, what is reflected by the local stability (21), itself a consequence of the fact that

$$-\varepsilon\left[\int_{\underline{\zeta}}^{h(2m)} f_{\zeta}(z)\,\mathrm{d}z\right]\kappa = 0,$$

independently of ε. In contrast, if network effects are absent, there is a negative drift on connectivity induced by any increase in volatility, a phenomenon captured in (12) by the term

$$-\varepsilon\left[\int_{\underline{\zeta}}^{v\frac{(1-\delta)}{\delta}} f_{\zeta}(z)\,\mathrm{d}z\right]\kappa < 0$$

for all $\varepsilon > 0$ and $\kappa > 0$. This makes it impossible that any sizable network density might be supported beyond a certain threshold on the rate of volatility. Thus, in combination with (10), a sharp discontinuous collapse of connectivity must occur when ε goes above that threshold.

6. Related literature

The approach pursued here bears on a number of different topics, not only in the fields of game theory and economics but also in sociology or the analysis of complex systems. Let me refer to each of them in turn.

First, we may view the theoretical framework proposed as similar in spirit to that of the evolutionary literature, where players are assumed to interact through a certain game and the long-run configuration is obtained through a gradual stochastic process of learning and adjustment. The early part of this literature (cf. Kandori et al. (1993), Young (1993), or Ellison (1993)) considered a set where the pattern of interaction of players (global or local) is fixed throughout, but recent work extends the analysis to a context where the network is not fixed but coevolves with players' game decisions – see e.g. Droste et al. (1999), Jackson and Watts (2002), or Goyal and Vega-Redondo (2004).

The latter approach, of course, bears a close relationship to the booming body of literature whose specific concern is the study of pure models of network formation. One of the earliest papers in this field was Aumann and Myerson (1989), with the more recent paper by Jackson and Wolinsky (1996) having played an important role in reviving interest in this topic. Whereas the approach of these papers is mostly static,[13] an explicit dynamic approach to the problem is undertaken in Bala and Goyal (2000).

[13] For example, Jackson and Wolinsky rely on a notion of pairwise-stability that is akin that introduced in Definition 1.

In the vast area of repeated games, there are two papers, Kandori (1992) and Ellison (1994), which share some motivation with our approach. They propose a model where a large population of players are repeatedly and randomly matched to play a Prisoner's Dilemma game. They find that, in this context, it may be still possible to induce cooperation through a social norm (equilibrium) that reacts to any deviation by punishing subsequent partners.[14] In a sense, the contrast here is analogous to that explained before concerning the early evolutionary literature with fixed and global interaction structure. That is, while the aforementioned papers embed players' interaction in a population context, they abstract from the effect of social *structure* on behavior by postulating a fixed and global pattern of play.

Still in the area of repeated games, three additional related papers are Bernheim and Whinston (1990), Haag and Lagunoff (2000), and Lippert and Spagnolo (2002). The first one studies a model of multimarket collusion where a group of firms participating in some common set of markets may decide to make their behavior in any one of them depend on what has been observed in other markets. Naturally, this enhances the collusion (i.e. cooperation) potential in ways analogous to those considered here. The key difference is that the flow of information is instantaneous and interaction is joint (but segmented), so that no phenomena arise analogous to those channeled here through the social network.

Instead, the paper by Haag and Lagunoff does study a setup where players are involved in repeated play of a Prisoner's Dilemma with partners specified by some *given* social network. Its approach, however, is mostly normative and static. An additional important difference is that players are always forced to play the same action – cooperate or defect – with everyone of their neighbors, so that the effect of the social network on behavior is made much more powerful than in the model studied in this paper.

Finally, in their independent paper, Lippert and Spagnolo introduce into a model of network-based RPD interaction a feature akin to that considered here, i.e. the assumption that information on what others did in the past moves only gradually through the network as players interact. Their setup is further enriched by allowing for the possibilities that interaction be directional (i.e. not necessarily reciprocal) and agents may control informational flows. They abstract, however, from the network formation issues that are our present main concern and focus instead on understanding how different simple network architectures bear on the extent of cooperation that can be supported at equilibrium.

The study of social networks has hardly been a preserve of economists or game theorists. Rather, it has long been a primary object of study by sociologists (cf. Granovetter (1973) and Burt (1982)) or applied psychologists (e.g. Milgram (1967)). The latter, for example, demonstrated through a clever simple experiment the surprisingly low number of steps which tend to separate any two arbitrarily chosen individuals in many large social networks. This phenomenon has recently attracted much attention by physicists and other researchers interested in the study of complex

[14] More recently, Bendor and Swistak (2001) have elaborated on this standpoint and provided a clear explanation of why in social contexts it might be often necessary to rely on "social norms that induce us to reward or punish people not for what they did to us but what they did to other members of one's group."

systems (cf. Watts and Strogatz (1998), Newman (2003), and Vega-Redondo (2005)). In part, their objective has been to propose natural network formation models – the so-called small-world contexts – that reconcile the aforementioned low distances among *randomly* selected individuals with another empirical regularity playing a key role in our model: high clustering (or, more generally, what we have called low *neighbor* distances).

Finally, I briefly discuss the literature on social capital, the notion that has been used in the Introduction to motivate the present approach. Rather than attempting a necessarily superficial survey of the vast and diverse range of research that goes under this heading, it should be more useful to focus on the work of Coleman (1988), arguably the author who (together with Putnam (1993)) brought the concept of social capital to prominence in socio-economic analysis. He is also one of the authors who has conceived it more in line with the view espoused here – see e.g. Coleman (1990, Chapter 12). For him, social capital is an inherently relational concept, to be regarded as an attribute of the social network. It is the key factor explaining the intensity and stability of socioeconomic interaction and also represents the basis of trust in repeated interaction (in particular, Coleman often uses the Prisoner's Dilemma as the paradigmatic setup to address this issue). However, for such a trust to emerge, what he calls the closure of the social network (analogous to our notion of cohesiveness) is generally key. He argues, moreover, that social capital is often underprovided, since the strong externalities associated with it are typically not internalized by individuals' own link-investment decisions. Finally, he stresses that social capital is a stock which, left to itself, depreciates with time and that, if it is to be (re)built successfully, must have inter-agent relations enjoy a sufficiently stable environment. The reader will recognize in these points many of the considerations that have informed (both in modelling and motivation) the analysis undertaken here. The empirical relevance of these considerations has also been confirmed in a wide number of contexts, ranging from international trade (Rauch (2001)), to credit (McMillan and Woodruff (1999)), insurance (de Weerdt (2002)), and the evolution of transition economies (Radaev (2002)). In all of these cases, cooperation has been found to crucially depend on the establishment of a social norm by which 'defectors' can be punished socially (cf. Footnote 14).

7. Summary and possible extensions

This paper has studied a model of network formation in which players are involved in an infinitely repeated Prisoner's Dilemma with each of her neighbors. Information on past behavior flows gradually along the network, a feature that impinges on the range of network configurations that can be supported as pairwise stable. The underlying payoff conditions change over time. This could sometimes deteriorate the stability of established links, but also creates the opportunity of forming new ones.

The analysis has focused on understanding how the interplay between the emerging (i.e. endogenous) characteristics of the network and its long-run performance depend on the key parameters of the model – most importantly, the *volatility rate*. By relying on the complementary use of numerical simulations and mean-field analysis, the main insights obtained may be summarized as follows. When players's *strategic* behavior is *network-based*, the architecture of the network adapts to increasing volatility by becoming more *cohesive*. This, in turn, acts as a buffer that mitigates the detrimental effects of volatility and allows to sustain a higher level of cooperative behavior (social capital). Instead, if *strategic* behavior is shaped in a strictly bilateral fashion (i.e. is *network-free*), the harmful effect of volatility is much more acute and its consequences materialize abruptly.

The model is quite stylized in a number of respects and therefore it could be fruitful to enrich it along various directions. Let me conclude with some suggestions.

A natural extension would involve enlarging the set of games under consideration, possibly considering other sorts of simple bilateral games (e.g. coordination games) or even 'playing the field' contexts. Along similar lines, one further option would be to explore the implication of assuming (as often done in the literature – recall Subsection 6) that each player must take the *same* action in *all* games she plays. The interrelation between the network considerations brought about by this important modification and the informational aspects studied here may add novel insights to existing models.

Concerning payoff volatility, it would be interesting to allow for the (arguably realistic) possibility that the realizations induced by any fresh payoff update may be correlated in some dimension. For example, it could be posited that the payoff draws enjoyed by a particular individual display positive correlation (as a reflection, say, of idiosyncratic characteristics) or, in a somewhat polar vein, that the draws obtained by different individuals include an aggregate component. Any of these modifications is bound to yield important implications on the network dynamics and its long-run architecture.

Finally, payoffs could be subject to some exogenous or endogenous trend. Concerning the first possibility (an exogenous trend), it would be interesting to understand the implications of letting payoffs be subject to some negative drift, as an additional manifestation of the 'Red-Queen Principle' mentioned above. As for the second possibility (an endogenous trend), it might be postulated that the payoffs earned by the different agents must be scaled (or normalized) by population-average payoffs or that new payoffs are drawn according to a moving distribution anchored to average (or frontier) conditions. In either case, this would add a supplementary competitive pressure on the agents and allow to conceive the process as one of economic growth.

Acknowledgements

I am grateful for useful comments by seminar participants at the Universities of Edinburgh, Essex, Queen Mary, Aix en Provence, the Institute for Advanced Studies at Vienna, the International Centre of Theoretical Physics at Trieste, and the

Fourgeaud-Roy Seminar at Paris. Finally, I want to acknowledge financial support by the Spanish MEC through the Grant No. SEJ 2004-02172.

Appendix

Proof of Proposition 1. Consider any link $ij \in g$ and focus, for concreteness, on player i. Pairwise-stability of this link requires that player i has incentives to cooperate with j under the threat that, if she were to do otherwise, the optimal penal code embodied by indefinite joint defection will be played by i and each of her neighbors $k \in N_i$ as soon as that deviation becomes common knowledge for both – an event that, in each of respective case, occurs $d^i(j,k)$ periods after the contemplated defection would take place. On the one hand, the payoffs following from a deviation with j are[15]:

$$\pi_i(Q^*(g) \overset{ij}{\setminus} D) = (1-\delta)(\zeta_{ij} + v) + \sum_{k \in N_i \setminus \{j\}} \left[\sum_{s=0}^{d^i(j,k)} (1-\delta)\delta^s \zeta_{ik} \right],$$

simply reflecting the fact that, for each $k \in N_i$, joint bilateral defection between i and k prevails from the period $t = d^i(j,k)$ onwards. On the other hand, if player i does not deviate with j, she obtains

$$\pi_i(Q^*(g)) = \sum_{k \in N_i} \zeta_{ik}.$$

Therefore, the stability condition

$$\pi_i(Q^*(g)) \geqslant \pi_i(Q^*(g) \overset{ij}{\setminus} D)$$

can be written as follows:

$$\delta \zeta_{ij} \geqslant (1-\delta)v - \sum_{k \in N_i \setminus \{j\}} \delta^{d^i(j,k)+1} \zeta_{ik}$$

which readily yields the stated condition. □

Proof of Proposition 2. Again consider any given link $ij \in g$ and focus on player i. In the present case, the payoff i obtains by unilaterally deviating with j is given by

$$\begin{aligned}\pi_i(Q^*(g) \overset{ij}{\setminus\setminus} D) &= (1-\delta)(\zeta_{ij} + v) \\ &\quad + \sum_{k \in N_i \setminus \{j\}} \left\{ \left[\sum_{s=0}^{d^i(j,k)-1} (1-\delta)\delta^s \zeta_{ik} \right] + (1-\delta)\delta^{d^i(j,k)}(\zeta_{ik} + v) \right\} \\ &= (1-\delta)(\zeta_{ij} + v) + \sum_{k \in N_i \setminus \{j\}} [(1-\delta^{d^i(j,k)+1})\zeta_{ik} + (1-\delta)\delta^{d^i(j,k)} v].\end{aligned}$$

[15] Recall that stage payoffs are normalized by the factor $(1-\delta)$.

Then, the equilibrium condition

$$\pi_i(Q^*(g)) = \sum_{k \in N_i} \zeta_{ik} \geqslant \pi_i(Q^*(g) \overset{ij}{\backslash\backslash} D)$$

can be written as follows:

$$\delta\zeta_{ij} + \sum_{k \in N_i \backslash \{j\}} \delta^{d^i(j,k)+1} \zeta_{ik} \geqslant (1-\delta)v + \sum_{k \in N_i \backslash \{j\}} (1-\delta)\delta^{d^i(j,k)} v$$

which is equivalent to (4). □

Proof of Proposition 3. The induced process is clearly aperiodic. Therefore, to establish the desired conclusion, it is enough to show that there is some particular state to which there is positive probability of returning, from any other state ω, in some finite number of steps. In the ensuing argument, the state that plays this recurrent role is the one where there are no links established between players, that is denoted by ω^e.

Consider any arbitrary state $\omega(t_0) = [(\vec{g}_i(t_0))_{i \in N}, (\zeta_{ij}(t_0))_{ij \in g}]$ prevailing at any time t_0. The first point to note is that, if n is large enough (and an even number, for simplicity), there is positive probability that the network might eventually be divided into two equal-sized subnetworks at some future $t_1 > t_0$, with no links across them. To see this, let N_1 and N_2 be two disjoint subsets of N of cardinality $n/2$. Suppose, without loss of generality, that $\zeta_{ij}(t_0) < \overline{\zeta}$ for all $ij \in g(t_0)$. Then, let revision events occur in sequence at $t > t_0$ so that players in N_1 obtain through global search payoff options ζ_{ij} with players $j \in N_1$ satisfying:

$$\overline{\zeta} > \zeta_{ij} > \max\left\{\frac{1-\delta}{\delta} v,\ \max_{k \in N_2} \zeta_{ik}(t_0)\right\} \tag{22}$$

and reciprocally for players in N_2. These revision opportunities induce pairwise-stable and payoff-improving links for each player in both N_1 and N_2 over those that can be supported with agents in the respective complementary sets. Therefore, if n is large enough (in particular, it is enough that $n \geqslant 2(2m+1)$, where m is the maximum number of links supportable by an agent), all links across N_1 and N_2 must eventually be removed at some $t_1 > t_0$.

But, clearly, if $n/2$ is sufficiently large (and again even, for simplicity), an analogous procedure separately operating within the aforementioned N_1 and N_2 may lead, at some $t_2 > t_1$, to the whole population getting divided into four equal-sized subnetworks with no links across them. This may then be repeated up to v times, with $v \geqslant \log(m+1)/\log 2$, so that at some t_v the population is divided into at least $m+1$ subnetworks with no links across them. For notational simplicity, relabel the set of nodes in each of these subnetworks as $N_1, N_2, \ldots, N_v$. The path leading to this outcome has positive probability, given the law of motion of the process.

Now suppose that a single player, say i, in one of those set of nodes, say N_q, is the only one to receive revision opportunities in the subsequent m periods, $t_v + 1, \ldots, t_v + m$, and these are of the global kind and pertain to players in m distinct N_r $(r \neq q)$. Further suppose that the entailed payoffs are high enough to induce player i to form the corresponding links – this has positive probability, in view of (22). Then, this player eventually removes all her preexisting links within N_q and all her existing links at $t_v + m$ connect her to m other groups N^r $(r \neq q)$. If these links are subsequently subject to payoff update and their payoffs are thus brought below $(1-\delta)/\delta v$, player i must have all her links removed and thus becomes an isolated player. Repeating this procedure with each of the remaining $n-1$ players in the population, the process may lead to the empty network at some $t \geqslant t_v + nm$, i.e. to the state ω^e. By construction, the chain of steps involved is finite and displays positive probability. The proof, therefore, is complete. □

Proof of Proposition 4. Consider the following two observations:

(1) Let $\eta > 0$ be such that $\overline{\zeta} - \eta > [(1-\delta)/\delta]v$. Then, since the density $f_\zeta(\cdot)$ is assumed continuous on its support $[\underline{\zeta}, \overline{\zeta}]$ and revision opportunities are independent across players and time, the following conclusion applies. For all $\theta \in (0,1)$, there is some T such that if $t \geqslant T$, there is probability no lower than $1-\theta$ that every player i has received (in preceding periods $\tau < t$) at least m link formation opportunities with distinct partners j and associated payoffs $\zeta_{ij} > \overline{\zeta} - \eta$.

(2) Any of the link opportunities described in (1) are pairwise stable. Therefore, choosing θ and T as above, there is probability no lower than $1-\theta$ that, if $t \geqslant T$, every player i is supporting m links at t (the maximum number), all of them with associated payoffs no lower than $\overline{\zeta} - \eta$.

Then, since η and θ in (1)–(2) can be chosen arbitrarily small, the desired conclusion readily follows. □

References

Abreu, D., 1998. On the theory of infinitely repeated games with discounting. Econometrica 56, 383–396.

Aumann, R., Myerson, R., 1989. Endogeneous formation of links between players and coalitions: an application of the Shapley value. In: Roth, A. (Ed.), The Shapley Value. Cambridge University Press, Cambridge.

Bala, V., Goyal, S., 2000. A non-cooperative model of network formation. Econometrica 68, 1181–1231.

Bendor, J., Swistak, P., 2001. The evolution of norms. American Journal of Sociology 106, 1445–1493.

Bernheim, D., Whinston, M., 1990. Multimarket contact and collusive behavior. Rand Journal of Economics 21, 1–26.

Burt, R.S., 1982. Towards a Structural Theory of Social Action. Academic Press, New York.

Carrol, L., 1872. Through the Looking Glass and What Alice Found There. Macmillan, London.

Coleman, J.S., 1988. Social capital in the creation of human capital. American Journal of Sociology 94, S95–S120.

Coleman, J.S., 1990. Foundations of Social Theory. Harvard University Press, Cambridge, MA.

de Weerdt, J., 2002. Risk-sharing and endogenous network formation. Discussion Paper, DP2002/57, WIDER, United Nations University.

Droste, E., Gilles, R., Johnson, K., 1999. Endogenous interaction and the evolution of conventions. Tilburg University, Mimeo.

Durlauf, S.N., 1999. The case 'against' social capital. University of Wisconsin, Mimeo.

Ellison, G., 1993. Learning, local interaction, and coordination. Econometrica 61, 1047–1071.
Ellison, G., 1994. Cooperation in the Prisoner's Dilemma with anonymous matching. Review of Economic Studies 61, 567–588.
Goyal, S., Vega-Redondo, F., 2004. Network formation and social coordination. Games and Economic Behavior 50, 178–207.
Granovetter, M.S., 1973. The strength of weak ties. American Journal of Sociology 78, 1360–1380.
Haag, M., Lagunoff, R., 2000. Social norms, local interaction, and neighborhood planning. International Economic Review, Forthcoming.
Jackson, M., Watts, A., 2002. On the formation of interaction networks in social coordination games. Games and Economic Behavior 41, 265–291.
Jackson, M., Wolinsky, A., 1996. A strategic model of economic and social networks. Journal of Economic Theory 71, 44–74.
Kandori, M., 1992. Social norms and community enforcement. Review of Economic Studies 59, 63–80.
Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.
Lippert, S., Spagnolo, G., 2002. Network of relations. University of Mannheim, Mimeo.
Marsili, M., Vega-Redondo, F., Slanina, F., 2004. The rise and fall of a networked society: a formal model. Proceedings of the National Academy of Sciences of USA 101, 1439–1442.
McMillan, J., Woodruff, C., 1999. Interfirm relationships and informal credit in Vietnam. Quarterly Journal of Economics 114, 1285–1320.
Milgram, S., 1967. The small-world problem. Psychology Today 2, 60–67.
Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Review 45, 167–256.
Neyman, A., 1985. Bounded complexity justifies cooperation in the finitely repeated Prisoner's Dilemma. Economic Letters 19, 227–229.
Putnam, R.D., 1993. Making Democracy Work: Civic Traditions in Modern Italy. Princeton University Press, Princeton.
Radaev, V., 2002. Enterpreneurial strategies and the structure of transaction costs in Russian business. Problems of Economic Transition 44, 57–84.
Rauch, J., 2001. Business and social networks in international trade. Journal of Economic Literature 39, 1177–1203.
Rubinstein, A., 1986. Finite automata play the repeated Prisoner's Dilemma. Journal of Economic Theory 39, 83–96.
Stanley, H.E., 1971. Introduction to Phase Transitions and Critical Phenomena. Oxford University Press, New York.
Van Valen, L., 1973. A new evolutionary theory. Evolutionary Theory, vol 1, 1–30.
Vega-Redondo, F., 2002. Building up social capital in a changing world. WP-AD 2002-26, Instituto Valenciano de Investigaciones Económicas.
Vega-Redondo, F., 2005. Diffusion, Play, and Search in Complex Social Networks. In: The Econometric Society Monograph Series, Cambridge University Press, Cambridge, Forthcoming.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.
Woolcock, M., 2000. Using Social Capital: Getting the Social Relations Right in the Theory and Practice of Economic Development. Princeton University Press, Princeton, NJ.
Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.