European University Institute

Department of Economics

EUI Working Papers

ECO 2008/08

Networks Emerging in a Volatile World

Gorge Ehrhardt, Matteo Marsili and Fernando Vega-Redondo

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

Networks Emerging in a Volatile World

GEORGE EHRHARDT, MATTEO MARSILI

and

FERNANDO VEGA-REDONDO

EUI Working Paper **ECO** 2008/08

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

Networks emerging in a volatile world

George Ehrhardt
The International Centre for Theoretical Physics, Trieste

Matteo Marsili
The International Centre for Theoretical Physics, Trieste

Fernando Vega-Redondo
European University Institute, Florence

First version: August, 2006
This version: November, 2007

Abstract

The paper proposes a model to study the conditions under which complex networks emerge (or not) when agents are involved in a dynamic coordination setup. The focus, however, is not on the classical issue of equilibrium selection – instead, our aim is to shed light on how agents' efforts to coordinate affect the process of network formation in a large and complex environment. The model posits that, over time, new links are created if they are profitable, and existing links disappear due to exogenous decay. Alongside this struggle between link creation and link destruction, agents' choices in the coordination game adapt to their current local conditions and thus coevolve with the social network. The dynamic behavior of the system is studied within different time scales (the long and ultra-long runs), which differ in the role accorded to the noise that is induced by finite populations. We characterize analytically the evolution of the model in each case and show that, depending on the time scale under consideration, the process displays discontinuous transitions in overall connectivity, resilient transformations in network topology, and equilibrium multiplicity. This behavior is akin to that observed concerning various network phenomena where coordination and network formation display mutually reinforcing roles.

JEL Classification nos.: C73, D83, D85.

1 Introduction

Social networks constitute the backbone underlying much of the interaction conducted in socioeconomic environments.[1] Therefore, when this interaction attains a *global* reach it must have, as its counterpart, the emergence of a social network with a wide range of overall (typically indirect) connectivity. Naturally, for such a social network to emerge, agents must be able to link profitably. But this in turn demands that they display a similar – at least compatible – behavior. Thus, for example, they must use coherent communication procedures, share key social conventions, or have similar technical ability. Here, we may quote the influential work of Castells (1996):

> Networks are open structures, able to expand without limits, integrating new nodes as long as they are able to communicate within the network, namely as long as they share the same communication codes (for example, values or performance goals).

Reciprocally, of course, such a convergence of behavior is facilitated by the range of social interaction being global rather than local or fragmented. This suggests the idea that the buildup of a global social network might be understood as the outcome of twin cross-reinforcing processes: one that facilitates the convergence of norms and behavior, and another that extends the range of social connectivity. This mutual reinforcement also suggests that if such a global transition indeed takes place, it should be relatively *fast* and *resilient*.

To explore these ideas, we propose a stylized model in which agents are involved in a local coordination game with their neighbors in a coevolving network. Two features characterize the dynamics. First, we postulate that while links vanish at a constant exogenous rate, new links are created only between agents who randomly meet and happen to be "coordinated," i.e. display the same action/strategy in the coordination game. On the other hand, we assume that, in the same time scale (i.e. at a comparable rate), agents adjust their action towards that which maximizes the extent of coordination with their current

[1] Some important examples include labor markets (Granovetter (1974), Montgomery (1991), or Calvó-Armengol and Jackson (2004)), informal insurance (Murgai *et al.* (2002) or Fafchamps and Lund (2003)), trade arrangements (Kirman *et al.* (2000) or Kranton and Minehart (2001)), the performance of firms and other organizations (Krackhardt and Hanson (1993), van Alstyne (1997), and Burt (1992)), industrial districts (Saxenian (1994) and Castilla et al. (2000)), and networks of interfirm collaboration in either research and development (Hagedoorn (2002) and Powell *et al.* (2005)) or joint ventures (Kogut *et al.* (2006)).

neighbors. For brevity, we refer to the first feature as *homophily* and the second as *conformity*.

Our abstract formulation can be regarded as an idealized representation of many network phenomena in the social world that embody some degree of homophily and conformity. To fix ideas, let us suppose that the population consists of a set of academics whose endeavor is to write papers on some wide area of research, say economics.[2] Each link represents ongoing (bilateral) collaboration, whose persistence depends on the fruitfulness of the research program in question – something that is *a priori* uncertain. Eventually, research programs become exhausted, which leads to the termination of the corresponding links. Thus, if the network of collaboration is to remain dense, researchers must keep creating fresh links with new partners. But, in so doing, the two parties need to exhibit some compatibility (coordination), which impinges on their ability to undertake fruitful collaboration.

The interplay between coordination games and networks has received ample attention in recent years. Initially, it was studied by Blume (1993), Ellison (1993), or Young (1998) in a context where players are involved in a fixed coordination game with their neighbors and the underlying network remains fixed. Subsequently, Jackson and Watts (2002), Ely (2003), and Goyal and Vega-Redondo (2005) have studied the issue in a setup where the network coevolves with agents' behavior. In every case the main concern has been to understand how the pattern of local interaction (evolving or not) impinges on equilibrium selection – typically, on the tension between risk dominance and efficiency.

In this paper, however, our primary concern is *not* one of equilibrium selection *per se* and thus we assume that the coordination game played by agents is extremely simple. In particular, all possible actions display a fully symmetric role, and hence the concern is not to single out the action on which the population may eventually coordinate. Rather, the focus is on the extent to which coordination (on whatever action) may be attained, in turn affecting the level (or density) of interaction that can be effectively supported. This, in essence, amounts to understanding how agents' struggle towards coordination may facilitate (or hinder) the build-up of the social network.

[2] Academic research collaboration has been recently studied in a variety of different disciplines, taking advantage of the availability of large-scale data readily available. For example, Newman (2001) has analyzed the fields of physics, biomedical research, and computer science; Grossman (2002) has focused on mathematics; and Goyal *et al.* (2006) has centered on economics.

We find that whether such a network build-up might indeed obtain in the long run is not a foregone conclusion – i.e. it depends on the underlying parameters and, possibly as well, on the initial conditions from which the process starts. But, when it indeed materializes, the resulting network structure is an intricate one, displaying substantial internode heterogeneities and no discernible patterns. This stands in marked contrast with the very simple architectures (e.g. involving completely connected components) that typically arise in the aforementioned equilibrium-selection literature. On the other hand, we also find that the process underlying the formation of the social network displays some interesting dynamic features:

(1) Sharp qualitative transitions occur, as a discontinuous response to slight changes in the underlying parameters.

(2) The system transitions in (1) display hysteresis, i.e. they are locally irreversible in the long run, even if the environmental parameters revert to their original values.

(3) The multiplicity resulting from (2) does not persist in the "ultralong run," where the ergodicity of the process yields a unique robust outcome.

In the dynamic model studied in this paper, the social network is in continuous flux, link creation and destruction ever changing the circumstances of individual nodes. This differentiates our approach from that pursued by much of the existing literature on social-network formation, which stresses the strategic dimension of the phenomenon and focuses on the equilibria (or absorbing points) of suitably defined network-formation games (or processes).[3] However insightful this approach indeed is for the study of small or/and quite stable scenarios, we believe that it is less appropriate for the study of large and volatile socioeconomic environments. For such contexts, a different methodology seems more in order, akin to that used by the modern theory of complex networks – see e.g. the so-called *small-world networks* introduced by Watts and Strogatz (1999) or the *scale-free networks* studied by Barabási and Albert (1999). This literature, however, typically relies on essentially *algorithmic* mechanisms of network construction that abstract from individual incentives or otherwise lack a social motivation.[4] Hence an additional motivation for the paper is

[3] See, for example, the seminal papers by Jackson and Wolinsky (1996) and Bala and Goyal (2000), or the exhaustive recent survey by Jackson (2005).

[4] See Newman (2003) or Vega-Redondo (2007) for complementary surveys of the literature. On the other hand, for another recent paper where this approach is applied to network for-

methodological, i.e. to illustrate that the statistical methods used by much of complex-network literature can be adapted to study the evolutionary processes of network formation underlying the social dynamics of large populations.

The rest of the paper is organized as follows. First, Section 2 presents the basic theoretical framework and describes in detail both the subprocesses of link formation and action adjustment. Then, in Section 3, we prove the ergodicity of the process, characterize its invariant distribution, and highlight the most prominent features of the unique configurations that, for large populations, absorb most of the weight in the "ultralong run" (i.e. in infinite time). In Section 4, we turn to studying the "long run," i.e. the shorter time scale where multiplicity is possible and initial conditions may play a decisive role. The main body of the paper concludes in Section 5 with a summary of our approach and a review of its main conclusions. For the sake of readability, the formal proofs of the results are gathered in an Appendix.

2 Theoretical framework

Let there be a certain population of agents, $P = \{1, 2, ..., N\}$, who interact bilaterally over time as specified by the evolving social network. Time is modelled continuously, with $t \in [0, \infty)$. At any t, the state of the system $\omega(t) = (\boldsymbol{\alpha}(t), G(t))$ consists of two items: an action profile and a network. The action profile $\boldsymbol{\alpha}(t) = (\alpha_i(t))_{i \in P} \in A^P$ specifies the action played by each agent, where $A = \{a_1, a_2, ..., a_q\}$ is the set of possible actions. The network, on the other hand, is described by an adjacency matrix $G(t) = (g_{ij}(t))_{i,j \in P}$, where an *undirected* link between i and j exists at t if $g_{ij}(t) \equiv g_{ji}(t) = 1$, while if this link does not exist then $g_{ij}(t) \equiv g_{ji}(t) = 0$. The set of all such possible states ω is denoted by Ω.

Players adjust both actions and links over time. The dynamics is described by a continuous time Markov process for the state $\omega(t)$, and is therefore completely determined by the rates governing all possible transitions $\omega \to \omega'$. These transitions pertain to adjustments that involve (a) link creation, (b) link destruction, (c) action revision. We now describe each of these in turn.

Link formation: At a certain positive rate η, each agent i receives a link creation opportunity. When such an opportunity arrives at some t, another

mation in social contexts, see Jackson and Rogers (2007). This paper contemplates a infinite sequence of entry decisions in which every new player searches for her partners according to some combination of random and network-based search.

agent j is randomly chosen in the population (all with the same probability). When no link exists between i and j (i.e. $g_{ij}(t) = 0$), this link is formed if, and only if, $\alpha_i(t) = \alpha_j(t)$. This rule embodies the referred idea of *homophily.* On the other hand, if the link ij is already part of the network (i.e. $g_{ij}(t) = 1$), it is supposed that the link remains in place.

Link removal: Existing links decay at a rate λ, which for simplicity is taken to be constant and exogenous. This component of the process may be conceived as capturing unmodelled environmental *volatility* that affects the value or feasibility – and thus the perdurability – of existing links.[5] Relatedly, it may also be viewed as a reflection of a capacity constraint in linking: since the probability that an agent maintains all her links decreases linearly with each additional one, the expected number of her prevailing links is finite.

Action revision: Finally, concerning action adjustment, we posit that, at every t, each agent i independently receives at a rate ν the opportunity to revise her current action. If such a revision opportunity materializes, the agent is assumed to choose one of the actions displayed by a plurality of her neighbors, possibly with some status quo bias probability $u \geq 0$ in case of ties. This is our embodiment of the notion of *conformity.* To formulate it precisely, let

$$B_i(t) \equiv \{a_r \in A : |\{\alpha_j(t) = a_r : g_{ij}(t) = 1\}| \geq |\{\alpha_j(t) = a_{r'} : g_{ij}(t) = 1\}| \ \forall r'\}$$

where $|\{\cdot\}|$ stands for the cardinality of the set in question, and denote by $I_i(t) \in \{0, 1\}$ the indicator function for the event $\alpha_i(t) \in B_i(t)$. Then, the probability $p_{ir}(t)$ that a revising agent i chooses any particular action a_r is

$$p_{ir}(t) = \begin{cases} \frac{1-I_i(t)u}{|B_i(t)|} & \text{if } a_r \in B_i(t) \backslash \{\alpha_i(t)\} \\ I_i(t)\left[\frac{1-u}{|B_i(t)|} + u\right] & \text{if } a_r = \alpha_i(t) \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Thus all actions are treated symmetrically, except for the possibility that (if $u > 0$) there might be some weak preference for the status-quo.

[5] One could generalize matters and (in line with the considerations motivating link creation) posit that links connecting nodes i, j such that $\alpha_i(t) \neq \alpha_j(t)$ decay at a possibly larger rate $\lambda' \geq \lambda$. However, as it will be clear from our analysis in subsequent sections, the (ultra)long-run behavior for this general formulation would always coincide with the particular case where $\lambda' = \lambda$. Thus, for the sake of simplicity, we focus on this scenario.

To fix ideas, the above adjustment framework may be conceived as one where myopic agents are involved in a pure-coordination game with their neighbors and links are costly to form. Specifically, suppose that the action space A defines the common strategy space of an underlying bilateral game with payoff function $\pi : A \times A \to \mathbb{R}$ given by:

$$\pi(a_r, a_{r'}) = \bar{\pi} \qquad \text{if } r = r' \tag{2}$$

$$\pi(a_r, a_{r'}) = 0 \qquad \text{if } r \neq r' \tag{3}$$

for some $\bar{\pi} > 0$. Then, under the assumption that each agent i chooses the same action $\alpha_i(t)$ in each of her bilateral encounters at t, her total gross payoff at that time is simply $\sum_{j:g_{ij}(t)=1} \pi(\alpha_i(t), \alpha_j(t))$. In addition, let us assume that the formation of any new link entails a cost $c(< \bar{\pi})$ for both agents involved so that, if the link ij is formed at some t, the *net* payoff of this *new* link for each agent is $\pi(\alpha_i(t), \alpha_j(t)) - c$.

In this coordination scenario, *homophily* is simply a consequence of the linking incentives faced by (myopic) agents when, given the prevailing actions, they are considering creating any new link. In line with the illustration contemplated in the Introduction – i.e. a set of academics who collaborate pairwise in their research – one possible interpretation of the different actions in A are alternative fields of specializations. Then, the implicit idea here is that the start of such a collaboration is fruitful (and thus justifies the linking cost) if, and only if, the two academics work in the same field of research.[6] On the other hand, the pressure towards *conformity* in this context merely derives from the fact that the payoff of any given action solely depends on the number of neighbors adopting it. Thus, given the links currently in place, the myopically optimal adjustment of an agent's action involves adopting any one currently in the plurality in her neighborhood.

At this point, it is worth stressing an important implicit assumption made by our model: only one link can be changed (created or destroyed) at any point

[6] As suggested by an anonymous referee, there is an even simpler version of the model that completely abstracts from action adjustment and would display equivalent behavior and analogous insights. Specifically, to continue with the illustration afforded by research collaboration, suppose that a researcher simply remains attached to a certain field as long as she has ongoing collaboration with some other coauthor working in it. Only an isolate researcher, that is, considers changing fields, perhaps out of disappointment of not finding viable collaborators. Since, as we shall see in Section 3 (cf. Proposition 1), our full-fledged model has the property that only isolate agents ever change their actions, the two formulations turn out to yield completely equivalent behavior.

in time, and agents cannot combine action and link adjustment. This contrasts with assumptions made by other papers of the network-formation literature, where actions and links can be adjusted simultaneously, and no stringent bound is imposed on the number of the latter that may be affected.[7] Even though this alternative formulation may be more appropriate for some applications (admittedly, the more realistic assumption would lie somewhere in between), a forceful implication of it will be to simplify drastically the network topologies that may endogenously arise. Thus, since the preservation of network complexity is one of our objectives here, we choose not to pursue that modelling route.

In sum, the overall dynamics of our model consists of three parallel subprocesses proceeding at corresponding different rates: action adjustment, at the rate ν; link formation, at the rate η; and link removal, at the rate λ. Without loss of generality, we shall normalize $\lambda = 1$ by simply scaling time appropriately. On the other hand, we shall find that neither the rate ν nor the probability u that govern action adjustment play any role in the characterization of the invariant distribution of the process.[8] This then leaves the rate η as the single remaining parameter on which we shall focus most of our analysis.

3 Asymptotic behavior

The first point to note is that the Markov process $\{\omega(t)\}$ is ergodic. This is the content of the following result, which also indicates that any state with links connecting agents choosing different actions is transient and thus is assigned a vanishing asymptotic probability by the invariant distribution.

Proposition 1 *Let* $\hat{\Omega} \equiv \{\omega = (\boldsymbol{\alpha}, G) \in \Omega : \forall i, j \in P,\ [g_{ij} = 1 \Rightarrow \alpha_i = \alpha_j]\}$. *The process* $\{\omega(t)\}$ *has a unique invariant distribution* μ *with* $\mu(\hat{\Omega}) = 1$.

Proof: See the Appendix.

The above result is a simple consequence of the fact that all links decay at a constant rate. Thus, the empty network can be reached from any state, and

[7] Consider, for example, the polar extreme formulation posited by Goyal and Vega-Redondo (2005) where agents can, at each point in time, both change their action in the underlying coordination game and create or destroy as many links as they desire. It is then easy to see that, in any stable configuration where the network is not empty, agents must be partitioned in at most two completely connected components (indeed, it often happens to be a single one).

[8] However, to simplify the stability analysis of the mean-field dynamics conducted in Section 4 we study the limiting case where $1 - u$ is infinitesimally low (but positive) and thus the bias in favor of the status in case of equal payoffs is large.

from an empty network only states in $\hat{\Omega}$ can be reached through subsequent link adjustment. This readily implies that, independently of initial conditions, a single recurrent class in $\hat{\Omega}$ must eventually absorb all paths of the stochastic process with probability one. Next, in Subsection 3.1, we characterize the invariant distribution for finite population size N and obtain some induced measures on interesting aggregates. Subsequently, in Subsection 3.2, we study the situation when the population size N grows large.

3.1 The invariant distribution

To characterize the invariant distribution established by Proposition 1, it is enough to identify *one* probability distribution on Ω that verifies the stationarity conditions embodying invariance. For, by ergodicity, there can be *only one* such probability distribution. Thus let $\rho(\omega \to \omega')$ denote the rate at which a transition from states ω to ω' occurs at any point in time. Then, it is clear that any probability distribution μ that satisfies the following conditions:

$$\sum_{\omega' \in \Omega} \mu(\omega')\rho(\omega' \to \omega) - \sum_{\omega' \in \Omega} \mu(\omega)\rho(\omega \to \omega') = 0 \qquad (\omega \in \Omega)$$

is stationary (or invariant) under the contemplated process. Building upon the above conditions, we find that the (unique) probability distribution that remains invariant is that specified in the following result.

Proposition 2 *The invariant probability measure μ is given by*

$$\mu(\omega) = \begin{cases} \Upsilon \prod_{i,j \in P,\, i<j} \left(\frac{2\eta}{N-1}\right)^{g_{ij}} & \text{if } \omega = (\alpha, G) \in \hat{\Omega} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\Upsilon \equiv \left[\sum_{\omega \in \hat{\Omega}} \prod_{i<j} (2\eta/(N-1))^{g_{ij}}\right]^{-1}$ is a normalizing constant (interpretable as the aggregate probability of all states associated to an empty network).

Proof: See the Appendix.

Having characterized through the invariant distribution μ the limit behavior of the process, we are now interested in "extracting" from it some of the interesting properties it imposes on the induced social structure. We shall begin with an assessment on the extent of social conformity entailed. Specifically, we want to know whether the population will settle in a dominant social norm (i.e. an action chosen by a plurality of the population) or, instead, no such social norm arises and all actions are equally represented in the population. Denote by n_r the fraction of agents choosing action a_r at some point in time and let $\mathbf{n} = (n_1, ..., n_q)$. Then, more precisely, our objective is to characterize the probability measure over such vectors $\mathbf{n}$ induced by the invariant distribution μ. To this end, we first introduce some useful notation.

Given any state $\omega = (\boldsymbol{\alpha}, G) \in \hat{\Omega}$, let $P_r(\omega) \equiv \{i \in N : \alpha_i = a_r\}$ be the set of agents choosing a_r, and let $N_r(\omega) \equiv |P_r(\omega)|$ be its cardinality. Correspondingly, denote by $\mathbf{P}(\omega) = (P_1(\omega), ..., P_q(\omega))$ and $\mathbf{N}(\omega) = (N_1(\omega), ..., N_q(\omega))$ the *profiles* associated to the particular state ω. Reciprocally, we also want to identify what states in $\hat{\Omega}$ are associated to any pair of specific such profiles, $\mathbf{P}$ and $\mathbf{N}$. That is, given any $\mathbf{P} = (P_1, ..., P_q)$ and $\mathbf{N} = (N_1, ..., N_q)$ (where $P_r \cap P_{r'} = \varnothing$ for all $r \neq r'$, $\bigcup_{r=1}^{q} P_r = P$, and $\sum_{r=1}^{q} N_r = N$) we are interested on the collection of states in $\hat{\Omega}$ that are consistent with them. These are respectively given by

$$\begin{aligned}
\hat{\Omega}(\mathbf{P}) &\equiv \{\omega = (\boldsymbol{\alpha}, G) \in \hat{\Omega} : P_r(\omega) = P_r \ (r = 1, ..., q)\} \\
\hat{\Omega}(\mathbf{N}) &\equiv \{\omega = (\boldsymbol{\alpha}, G) \in \hat{\Omega} : |P_r(\omega)| = N_r \ (r = 1, ..., q)\}.
\end{aligned}$$

Our aim is to have a precise assessment of the probabilities associated to the different profiles $\mathbf{N}$ that specify how many agents (rather than their specific identity) display each of the q possible actions. Thus we want to compute $\mu(\hat{\Omega}(\mathbf{N}))$ for each possible $\mathbf{N}$. To do so, we need to add the probability associated to all possible partitions $\mathbf{P}$ that are consistent with any given vector $\mathbf{N}$. A useful preliminary observation in that respect is that, for every $\omega \in \hat{\Omega}$, we can simply write:

$$\mu(\omega) = \Upsilon \prod_{r=1}^{q} \prod_{i,j \in P_r(\omega),\, i<j} \left(\frac{2\eta}{N-1} \right)^{g_{ij}}.$$

Next, we compute, for any given $\mathbf{P}$,

$$\begin{aligned}
\mu(\hat{\Omega}(\mathbf{P})) &= \sum_{\omega \in \hat{\Omega} : \mathbf{P}(\omega) = \mathbf{P}} \mu(\omega) \\
&= \Upsilon \prod_{r=1}^{q} \left\{ \prod_{i,j \in P_r(\omega),\ i<j} \left[\sum_{g_{ij}=0,1} \left(\frac{2\eta}{N-1} \right)^{g_{ij}} \right] \right\} \\
&= \Upsilon \prod_{r=1}^{q} \left\{ \left(1 + \frac{2\eta}{N-1} \right)^{\frac{1}{2}|P_r|(|P_r|-1)} \right\}.
\end{aligned}$$

And, finally, adding over all partitions $\mathbf{P}$ that are consistent with each given $\mathbf{N}$, we obtain:

$$\begin{aligned} \mu(\hat{\Omega}(\mathbf{N})) &= \Upsilon \sum_{\{P:\ |P_r|=N_r\ (r=1,..,q)\}} \left\{ \prod_{r=1}^{q} \left[\left(1+\frac{2\eta}{N-1}\right)^{\frac{1}{2}N_r(N_r-1)} \right] \right\} \\ &= \Upsilon \frac{N!}{\prod_{r=1}^{q} N_r!} \prod_{r=1}^{q} \left[\left(1+\frac{2\eta}{N-1}\right)^{\frac{1}{2}N_r(N_r-1)} \right]. \end{aligned} \tag{5}$$

The previous expression represents a key starting point of our analysis. It defines the probability measure ζ over the sizes of *action classes* – i.e. set of nodes displaying each possible action – given by $\zeta(\mathbf{N}) = \mu(\hat{\Omega}(\mathbf{N}))$. As it turns out, the induced measure $\boldsymbol{\zeta}$ underlies most of the asymptotic properties of interest of our process. In particular, it provides exhaustive information on network connectivity. To verify this point simply note that, by virtue of the factorized form (4) of the invariant distribution, conditional on any action profile $\mathbf{N} = (N_1, ..., N_q)$, one can write:

$$P\{k_i = k | \alpha_i = a_r\} = \binom{N_r - 1}{k} \left(\frac{2\eta}{N-1}\right)^k \left(1 - \frac{2\eta}{N-1}\right)^{N_r-1-k} \tag{6}$$

for every $r = 1, 2, ..., q$. (Here, of course, we use again the fact that no links eventually persist between agents displaying different actions.) Thus the conditional probability distributions over degrees prevailing within every action class are simply binomial and, as advanced, solely depends on the class-size distribution $\boldsymbol{\zeta}$. In the next subsection, we want study the form of this distribution when the population is large. In this case, matters are simplified substantially and sharper predictions can be obtained.

3.2 Large populations

A first straightforward observation is that the degree frequencies among nodes choosing a particular action a_r is Poisson distributed when $N \to \infty$, with an expected (and average) degree proportional to the fraction $n_r \equiv N_r/N$ of nodes choosing that action. This simply follows from the implied stochastic independence across links and the standard observation that the Binomial expression (6) converges to the Poisson distribution as N grows unboundedly and each N_r/N converges to some given n_r. Thus we find, in sum, that the statistics of network topology within each class is that of Erdös-Renyi random graphs. For convenience, we gather this conclusion in the following result.

Proposition 3 *For large N, the frequency of agents $i \in P_r$ who have k neighbors is arbitrarily well approximated by*

$$P\{k_i = k | \alpha_i = a_r\} = \frac{(2\eta n_r)^k}{k!} e^{-2\eta n_r} \tag{7}$$

where $n_r = N_r/N$ is the limiting fraction of agents in class r and $2\eta n_r$ their average degree.

Let us now turn to studying the induced distribution of component sizes, which as explained above represents the cornerstone of the analysis. Our essential result in this respect is that, for large N, the distribution (5) is arbitrarily well approximated by an expression of the form

$$\mu(\hat{\Omega}(\mathbf{N})) \simeq e^{-Nf(\mathbf{n})} \tag{8}$$

for some function $f : \Delta^{q-1} \to \mathbb{R}$ of $\mathbf{n} = (n_1, n_2, ..., n_q)$, where $n_r \equiv N_r/N$ is the fraction of agents adopting action a_r and Δ^{q-1} is the $(q-1)$-dimensional simplex. Note that the above expression sharply decouples the effect of population size N from that channeled through class frequencies $\mathbf{n}$. This will allow us to conduct an effective analysis of the limit scenario where N is taken to grow arbitrarily large. The following result states this conclusion more precisely, and also provides a specific form for the function f.

Proposition 4 *Given any $\mathbf{n}$, and any sequence of $\mathbf{N}$ such that $\mathbf{N}/N \to \mathbf{n}$ as $N \to \infty$, the limit invariant probability measure satisfies*

$$\begin{aligned} -\lim_{N\to\infty} \frac{1}{N} \log \mu(\hat{\Omega}(\mathbf{N})) &= f(\mathbf{n}) \\ &= f_0 + \sum_{r=1}^{q} \left[n_r \log n_r - \eta n_r^2 \right] \end{aligned} \tag{9}$$

for some constant f_0.

Proof: The proof – an iterated application of Stirling's formula – can be found in the Appendix.

The expression (8) that approximates the invariant distribution for large N indicates that, in this limit, we are particularly interested in the minima of the function f. Indeed, these minima correspond to the set of configurations (i.e. class-size distributions) around which the process will concentrate most of its asymptotic probability – or spend most of its time as $t \to \infty$.

Consider, therefore, the following minimization problem:

$$\min_{\mathbf{n} \geq \mathbf{0}} f(n_1, n_2, ..., n_q) \tag{10}$$

$$\text{s.t.} \quad \sum_{r=1}^{q} n_r = 1. \tag{11}$$

As a first step in finding the solution to such a constrained optimization problem, it is useful to focus on the following necessary First-Order Conditions (FOC):

$$n_r \, e^{-2\eta n_r} = e^{\beta - 1} \qquad (r = 1, ..., q) \tag{12}$$

where β is a Lagrange multiplier, determined so as to enforce the normalizing constraint (11). Of course, the above conditions (in combination with suitable second-order conditions, as explained below) are only able to single out local minima of f, not necessarily global ones. Local minima, however, are of interest for two reasons. First, they are the sole candidates to minimize f globally. Second, they will be seen to be of interest in themselves, as relevant predictions in a time scale shorter than that associated to ergodicity.

In view of the properties of the function $n_r e^{-2\eta n_r}$,[9] a stark conclusion readily follows from (12): there are at most two possible values for each n_r – which we shall denote by n_+ and n_- ($n_+ \geq \frac{1}{2\eta} \geq n_-$) – that can be part of a solution to the above optimization problem. This allows one to divide the set of actions into two categories alone: the relatively predominant ones – whose *common* frequency is n_+ – and the ones in relative minority – whose frequency is n_-. *A priori*, the *number* of actions associated to either of the two situations is indeterminate, so we shall denote by L_+ and $L_- = q - L_-$ their respective numbers in the solution to be found.

A direct implication of (12) is that $n_+ e^{-2\eta n_+} = n_- e^{-2\eta n_-}$. This, in combination with the normalization condition $L_+ n_+ + (q - L_+) n_-$, allows one to determine uniquely the values of n_+ (and n_-) for all L_+ and η. With a little algebra we find that n_+ is implicitly defined by the solution of the following

[9] The function $h(x) = xe^{-2\eta x}$ has a unique maximum at $x = \frac{1}{2\eta}$, and $\lim_{x \to \infty} h(x) = h(0) = 0$.

equation:

$$n_+ = \left[L_+ + (q - L_+) e^{-2\eta \frac{qn_+ - 1}{q - L_+}} \right]^{-1} \tag{13}$$

whereas $n_- = (1 - L_+ n_+)/(q - L_+)$ is given by normalization.

The value of L_+ is one of the important features singling out a solution to the system of FOC (12). Another important consideration is, of course, whether it also satisfies the relevant second-order conditions of the optimization problem. Only then can one guarantee that it is a local minimum of f and thus a candidate for global minimum. Naturally, we find that the implications of such a second-order requirement as well as the range of possible values of L_+ displayed by a solution of the FOC crucially depends on the prevailing parameters of model: q and, most crucially, η. A full characterization of the situation is provided by the following result.

Proposition 5 *Given $q \geq 2$, the pattern of local minima of the function f has, as a function of η, a structure characterized by two thresholds, $\check{\eta}$ and $\hat{\eta}$ with $\check{\eta} \leq \hat{\eta} \equiv q/2$, and the following structure:*

(i) There exists an (action-symmetric) local minimum with $L_+ = 0$ if, and only if, $\eta \leq \check{\eta}$. This minimum is unique in the class with $L_+ = 0$.

(ii) There exist a collection of q local minima with $L_+ = 1$ – one for each action a_r being the dominant action – if, and only if, $\eta \geq \hat{\eta}$. These q (action-asymmetric) minima are unique in the class with $L_+ = 1$ and have n_+ increase with η.

(iii) There are no minima with $L_+ \geq 2$.

Proof: See the Appendix.

The solution n_+ of equation (13) is shown in Fig. 1 as a function of η for $L_+ = 1$. The condition of stability – i.e. that n_+ increases with η – readily allows us to distinguish minima (solid line) from saddle points (dashed line). The $L_+ = 0$ solution $n_r = 1/q$ is also shown. Overall, we can distinguish three different ranges for the parameter η in which qualitatively different situations are possible. First, for $\eta \in (0, \check{\eta})$, f has only the symmetric ($L_+ = 0$) minimum. This corresponds to a configuration $\mathbf{n}$ where all q actions display the same frequency $n_r = n_- = 1/q \leq \frac{1}{2\eta}$. Second, for $\eta \in (\check{\eta}, \hat{\eta})$, two qualitative different possibilities arise. On the one hand, a symmetric configuration just as before is possible. But, on the other hand, there are also q asymmetric configurations ($L_+ = 1$) where a single action a_r enjoys a high frequency $n_r = n_+ > 1/q$ and the other $q-1$ actions $a_{r'}$ ($r' \neq r$) display a relatively low one $n_{r'} = n_- < 1/q$. Finally, for the third range where $\eta \in (\hat{\eta}, \infty)$, only asymmetric configurations of the latter sort exist. Apart from its asymmetry in actions' frequencies, the $L_+ = 1$ solutions differ from the $L_+ = 0$ one in that, rather than having the network being highly fragmented in small components, its larger action class is characterized by a giant component, which connects a significant (i.e. nonvanishing) fraction of all nodes. This is a consequence of the following facts: *(i)* within each component, the network is an Erdös-Renyi random graph (Proposition 3) with average degree $2\eta n_r$; *(ii)* the values for n_- and n_+ satisfy $2\eta n_- < 1 < 2\eta n_+$; *(iii)* a giant component exists almost surely in Erdös-Renyi random graphs if, and only if, the average degree is larger than one.

Thus, when either $\eta < \check{\eta}$ or $\eta > \hat{\eta}$, Proposition 5 provides a unique prediction for the asymptotic configurations $\mathbf{n}$ induced by the process as $t \to \infty$. For the intermediate range $\eta \in (\check{\eta}, \hat{\eta})$, the issue of which of the two alternative configurations (with $L_+ = 0, 1$) eventually prevails hinges upon the sign of $\Lambda(\eta) \equiv f(1/q, ..., 1/q) - f(n_+(\eta), , ..., n_-(\eta))$. In fact, as it might be expected, we find that there is a single *threshold value* $\eta^* \in (\check{\eta}, \hat{\eta})$ such that, for all $\eta \geq 0$, we have:

$$\Lambda(\eta) \geq 0 \Leftrightarrow \eta \geq \eta^*. \tag{14}$$

One can readily check by direct substitution that the precise value of the threshold is given by $\eta^* = [(q-1)/(q-2)] \log(q-1)$. Indeed, from (13), we may first obtain that, at $\eta = \eta^*$, the corresponding value for the size of the large class is $n_+ = 1 - 1/q$ and, therefore, from normalization, $n_- = 1/[q(q-1)]$. Then, if one now introduces those values in $\Lambda(\cdot)$, it readily follows that $\Delta(\eta^*) = 0$, as required by the threshold condition (14). The value of η^* for $q = 10$ is marked in Figure 1 by a dashed vertical line. Figure 2 below, on the other hand, depicts how η^* varies with q, as this threshold is being bracketed by the two values – $\check{\eta}$ and $\hat{\eta}$ (also functions of q) – that define the boundaries of the coexistence region.

As explained, the fact that, in the infinite-time limit, the process predicts a unique configuration (i.e. class-size distribution) is a direct consequence of the expression obtained for the invariant distribution for large N in terms of the function f – generically, this function must display a unique *global* minimum (or

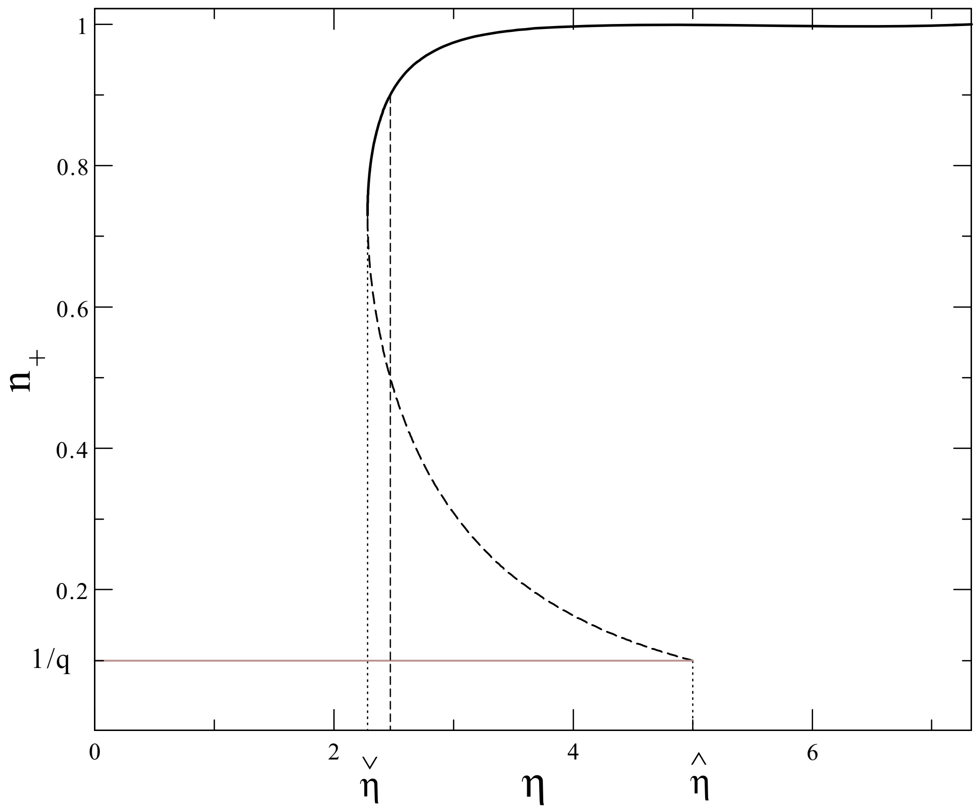

Figure 1: Solution of Eq. (13) with $L_+ = 1$ as a function of η, for $q = 10$. The solid line corresponds to a minimum of f, wheras the dashed line corresponds to a saddle point. The horizontal line is the solution in the $L_+ = 0$ case: $n_r = 1/q$, which exists for $\eta < q/2$.

a collection of isomorphic ones).[10] But in the parameter region where $\eta \in (\check{\eta}, \hat{\eta})$ such a global minimum coexists with another local minimum that, intuitively, should also tend to strongly absorb the configurations in its vicinity when N is very large. This suggests the consideration of two different time scales in which the predictions for the model can be alternatively formulated. Specifically, following Binmore, Samuelson, and Vaughan (1995), let us differentiate between

[10] As has been explained, when $\eta > \eta^*$, the global minimum of f is in fact not unique and the overall probability mass is uniformly divided by the invariant distribution among q alternative configurations – one for each action as the dominant norm. Since these q configurations are isomorphic (only the identity of the dominant action is changed), we can naturally think of them as essentially capturing a unique prediction.

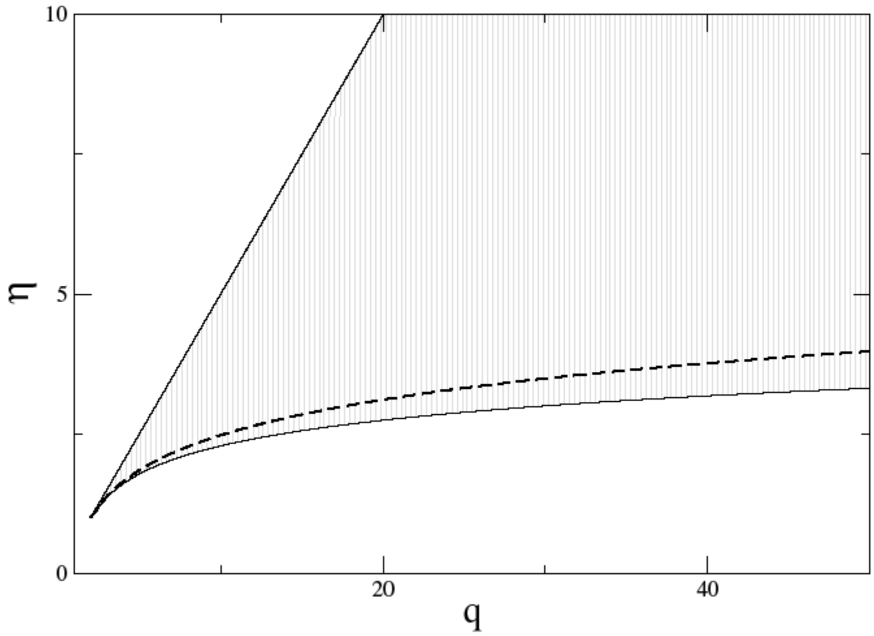

Figure 2: The value of η^* as a function of q (given by the dashed curve), which lies between the values of $\check{\eta}$ (lower solid curve) and $\hat{\eta}$ (upper solid curve) that are specified in Proposition 5.

the *long run* and the *ultralong run*. The latter is the prediction that builds upon the ergodicity of the process and is willing to wait an unbounded length of time before the mass asymmetries assigned to different configurations by the invariant distribution have had enough time to materialize, independently of initial conditions. This is indeed the gist of the analysis carried out so far, which singles out a uniquely predicted configuration (depending on how the parameter η lies relative to the threshold).[11]

In contrast, the long run is to be conceived as a time scale in which initial conditions may matter. Heuristically, it seems clear that if N is large and the system starts at a configuration that is close to a *local* minimum of f (i.e. a local maximum of the invariant distribution), steep "probability gradients" around

[11] As a follow-up of the discussion contained in Footnote 10, it is clear that, for $\eta > \eta^*$, one can conceive of a time scale even beyond the ultralong run in which no single action stands out from the rest, and the system undergoes transitions across different global minima of the function f. This regime, however, involves time scales that are so large that we choose to ignore them here.

it when N is large will bring the system towards that minimum with very high probability. In fact, it accords with the same intuition that the process should stay within a small neighborhood of that configuration for a long period of time (again, if the population is indeed large). This state of affairs, to be sure, can only last a finite amount of time (and also represent an infinitesimal fraction of the total as $t \to \infty$) if the configuration in question does *not* define a global minimum of f. We would be entitled, however, to speak of it as a "long-run" configuration because it could last for a very long time (in expected terms) if the population is sufficiently large. As it turns out, the informal description of the situation just outlined is essentially correct and can be made precise in a number of different ways. In the next section, we choose to do so through the route afforded by the so-called mean-field analysis of the stochastic process.

4 The long run: a mean-field approach

As explained, we want to identify the long run behavior of the process as that which obtains (with high probability) for large but finite times. Formally, such long-run behavior may be accessed by taking *first* the limit of infinite population size ($N \to \infty$) and then focusing on the large-time behavior ($t \to \infty$). In this approach, the dynamics of population frequencies is taken to follow a deterministic continuum-time dynamics – the Mean-Field Dynamics (MFD). In mathematical terms, the MFD is simply given by a system of ordinary differential equations that embodies *expected motion.* Intuitively, such mean-field analysis downplays the effect of the noise induced by a finite population, i.e. it implicitly assumes that within *any finite* time horizon T under consideration, the population is so large that any finite-sytem randomness can be safely ignored.

In contrast, the ergodic analysis of the preceding section reflects a reciprocal limit order: first $t \to \infty$, then $N \to \infty$. In this case, the intuition is that, given any population size, one is ready to wait an unbounded amount of time for any dependence on initial conditions to vanish completely. The fact that those limit operators do not always commute is, in a nutshell, what underlies the important conceptual dichotomy between the long and the ultralong run.

4.1 Mean-field dynamics: formulation

The mean-field dynamics is defined on an augmented description of the population frequencies considered in Section 3. Relying, for simplicity, on analogous notation, the states of the system consist of the current vector of population *frequencies* $\mathbf{n} \equiv [(n_{r,k})_{k=0}^{\infty}]_{r=1}^{q}$ specifying the frequency $n_{r,k}$ of agents in the population that choose each possible action r and have every possible degree k. The set of all such vectors – also called *population states* – will be denoted by Φ. Then, the *mean-field dynamics* (MFD) is defined as the continuous-time ODE on Φ whose induced paths satisfy, for each r and k, the equation

$$\dot{n}_{r,k} = \lim_{\Delta t \to 0} \frac{E[\Delta n_{r,k}]}{\Delta t} \equiv F_{r,k}(\mathbf{n}) \tag{15}$$

where $E[\Delta n_{r,k}]$ represents the expected change in $n_{r,k}$ in any time interval of infinitesimal size Δt. The exact form of this dynamics is provided by the following result.

Proposition 6 *The vector field $F(\mathbf{n}) \equiv [F_{r,k}(\mathbf{n})]_{r,k}$ of the mean-field dynamics is given by:*

$$F_{r,k}(\mathbf{n}) = 2\eta n_r n_{r,k-1} - 2\eta n_r n_{r,k} + (k+1) n_{r,k+1} - k n_{r,k}, \quad k > 0 \tag{16}$$

$$F_{r,0}(\mathbf{n}) = -2\eta n_r n_{r,0} + n_{r,1} + (1-u)\nu \sum_{s=1}^{q} [n_{s,0} - n_{r,0}] \tag{17}$$

where

$$n_r \equiv \sum_k n_{r,k}. \tag{18}$$

Proof: See the Appendix.

While the detailed derivation of the different terms in each $F_{r,k}$ is given in the Appendix, they can be succinctly explained as follows. The terms in (16) proportional to η describe the effect of the link creation subprocess on the population density, whereas all the other terms ("proportional" to $\lambda = 1$) describe the link decay subprocess. Either subprocess can increase or decrease each $n_{r,k}$, which is why each of them is responsible for two terms with opposite sign. The nodes with degree $k = 0$ (i.e. isolates) are special both because some of the terms of the equation for $k > 0$ do not occur (there are no nodes with degree $k = -1$), and because it is also affected by the exchange subprocess across different action classes $r \leftrightarrow s$. This subprocess occurs at the rate $\tilde{\nu} \equiv (1-u)\nu$ at which a node receives an action revision opportunity *and* is free of status-quo preference (i.e. decides to change actions despite the fact that there can be no payoff consequence of doing so).

It is intuitive that the MFD should represent a good (probabilistic) approximation of the behavior of the stochastic process for large N. Indeed, this can be rigorously confirmed by an adaptation of existing results from the modern literature on so-called stochastic approximation theory – see, specifically, the work by Benaïm and Weibull (2003*a*), which focuses on stochastic evolutionary dynamics.[12] To be precise, there are two complementary ways in which the MFD (15) approximates the original stochastic process:

1. Let $\{\mathbf{x}(t) \in \Phi\}_{t\geq 0}$ be the stochastic process induced by the model on the set of population states. Let $\{\mathbf{n}(t, \mathbf{x}_0)\}_{t\geq 0}$ be a solution of the ODE (15) with initial conditions $\mathbf{n}(0, \mathbf{x}_0) = \mathbf{x}_0$. Then, for any time horizon T, and any given $\varepsilon, \delta > 0$, there exists some lower bound $\hat{N}$ on population size such that if $N \geq \hat{N}$, then $\Pr\left[\max_{0\leq t\leq T} \|\mathbf{x}(t) - \mathbf{n}(t, \mathbf{x}(0))\| \geq \varepsilon\right] \leq \delta$, where $\|\cdot\|$ stands for the sup norm.

2. Let $\mathbf{n}^*$ be an asymptotically stable state of the ODE (15). There exists some U, an open (Borel) subset of Φ that includes $\mathbf{n}^*$, such that for any time horizon T, and any given $\varepsilon > 0$, there is some lower bound $\hat{N}$ on population size such that, if $N \geq \hat{N}$, the random variable $\tau(U) \equiv \inf\{t \geq 0 : \mathbf{x}(0) \in U,\ \mathbf{x}(t) \notin U\}$ giving the first exit time from U satisfies $\Pr\left[\tau(U) \geq T\right] \geq 1 - \varepsilon$.

The preceding statements specify two related forms in which it can be formally argued that the mean-field dynamics is a good approximation of the stochastic process in the long run. Verbally, the first one asserts that the population

[12] See, in particular, Section 6 of Benaïm and Weibull (2003*a*), in which they study stochastic processes where, as in the present case, the arrival of adjustment opportunities is governed by independent Poisson clocks. The essential feature shared by our framework and theirs is that, at every point an adjustment event takes place, it can only involve (with full probability) a *bounded* change in the characteristics of a *finite* number of nodes. This allows one to reproduce all the essential steps in their analysis, except for a minor and inconsequential adaptations. Let us, however, discuss two specific technical points of some importance.

First, the framework originally studied by Benaïm and Weibull presumed that the approximating vector field (i.e. the corresponding transition probabilities) can be defined independently of population size. But in many applications (the present one being an example), this is not the case. This issue has been addressed in a later paper by these same authors (Benaïm and Weibull (2003*b*)), where they generalize their previous analysis in this direction.

Second, those authors presume that the vector field is finite-dimensional. In our case, this would require bounding the support of the degree distribution by some $\bar{k}$, arbitrarily large but finite. This could be done, for example, by considering a variant of the model where no agent can support more than $\bar{k}$ links, so that if a linking opportunity arrives to an agent with that many links it cannot be materialized. In practice, however, it is easy to see that our model (in particular, as it pertains the stationarity and stability of its equilibria) would be continuous in such a parameter $\bar{k}$ at infinity. For simplicity, therefore, we choose to maintain our assumption throughout that $\bar{k} = \infty$, thus dispensing with a parameter of minor interest.

path induced by the stochastic process and that resulting from the Mean-Field Dynamics (MFD) are arbitrarily close, for an arbitrarily long period of time and with an arbitrarily high probability, provided the population is large enough. The second statement, on the other hand, focuses on asymptotically stable states of the MFD and indicates that any one of them is a robust long-run prediction in the following sense: whenever the process starts close to it, the ensuing path is very likely to remain also close for a very long period of time if the population is large.[13] Both conclusions, 1 and 2, allow us to view the MFD as a suitable description of the process in the limit $N \to \infty$, so that even for large but finite times (i.e. the long run) the "ergodicity-inducing" noise displayed by the finite-population process can be largely ignored.

4.2 Mean-field dynamics: analysis

Motivated by the preceding discussion, the objective of this subsection is to characterize the dynamic paths of (15). To simplify matters, we choose to focus on the natural case where, in case of payoff equivalence, the agents display a strong preference for the status quo – i.e. we make $u \nearrow 1$ and thus $\tilde{\nu} \equiv (1-u)\nu$ becomes infinitesimally small but positive. As we shall explain below, numerical simulations indicate that this simplifying assumption does not affect the predictions of the model.

First, notice by directly taking the sum of (15) over $k \geq 0$ that the evolution of each aggregate frequency n_r is given by

$$\dot{n}_r = \tilde{\nu} \sum_{s=1}^{q} \left[n_{s,0} - n_{r,0} \right], \tag{19}$$

which is proportional to $\tilde{\nu}$ and thus becomes very slow in this context. In contrast, the dynamics of each constituent $n_{r,k}$ in (16)-(17) has terms which do not vanish as $\tilde{\nu} \to 0$. And, moreover, if one neglects terms proportional to $\tilde{\nu}$ in (16)-(17), populations evolve in an independent fashion within each action class $r = 1, \ldots, q$.

[13] Indeed, much stronger conclusions can be derived, as discussed in detail by Benaïm and Weibull (2003*a*). For example, they find that the value of δ in the first conclusion can be made to decrease exponentially in population size. Or, concerning the second conclusion, they also prove (by a suitable application of the Borell-Cantelli Lemma) that for any arbitrarily large T, it is almost sure that for all but finitely many values of N we have $\tau(U) < T$, where U is a sufficiently small neighborhood of an asymptotically stable state $\mathbf{n}^*$ – or, equivalently put, that $\Pr\left[\lim_{N\to} \tau(U) = +\infty\right] = 1$, where note that $\tau(U)$ is to be conceived, in this case, as a function of N.

The former considerations suggest that, in order to study the context with small $\tilde{\nu}$, it is helpful to introduce a slow time variable

$$\tau = \tilde{\nu} t \tag{20}$$

such that, in the limit $\tilde{\nu} \to 0$, the dynamics of link adjustment in each class r is decoupled from the dynamics of the class sizes n_r. This happens because, in this limit, the two dynamics effectively operate in two different time scales: the former infinitely faster that the latter. In the next two subsections, we study each of these dynamics in turn: first, that of the network for given (slow) evolution of class sizes; second, the (fast) dynamics of class sizes.

4.3 Network dynamics

In this section, we posit that class sizes evolve slowly as specified by some given functions $n_r(t)$, and then find a solution of the ODE (15) under the conditions:

$$n_{r,k}(t=0) = n_{r,k}^{(0)} \quad \forall\, k \geq 0 \tag{21}$$

$$\sum_{k\geq 0} n_{r,k}(t) = n_r(t). \tag{22}$$

To carry out the analysis, it is convenient to describe the degree distribution prevailing within each class r through its Poisson representation given by

$$n_{r,k}(t) = \int_0^\infty dx \frac{x^k}{k!} e^{-x} f_r(x,t), \qquad \forall k \geq 0, \qquad (r = 1, \ldots, q) \tag{23}$$

where f_r is some suitable integrable function.[14] This representation is particularly useful because it turns a system of infinitely many coupled ODE's into a

[14] A constructive way to derive $f_r(x,t)$ from $n_{r,k}(t)$ is through its expansion in Laguerre polynomials

$$f_r(x,t) = \sum_{n=0}^{\infty} c_{n,r}(t) L_n(x) \tag{24}$$

where

$$L_n(x) = \sum_{k=0}^{n} \ell_{n,k} x^k = \sum_{k=0}^{n} \binom{n}{k} \frac{(-x)^k}{k!}$$

is the n^{th} Laguerre polynomial. In order to derive $c_{n,r}(t)$, multiply Eq. (23) by $k!\ell_{n,k}$ and sum over $k = 0, \ldots, n$. Using Eq. (24) and the orthonormality of Laguerre polynomials, we find

$$c_{n,r}(t) = \sum_{k=0}^{n} k!\ell_{k,n} n_{k,r}(t)$$

which is the desired expression.

single partial differential equation, as indicated in the following result.

Proposition 7 *A solution of the system of ODE (15) is given by (23) for each* $r = 1, ..., q$ *where* $f_r(x,t)$ *is a solution of the partial differential equation (PDE)*

$$\frac{\partial}{\partial t} f_r = \frac{\partial}{\partial x} \left[(x - 2\eta n_r) f_r \right] \tag{25}$$

with initial condition $f_r(x, 0) = \phi_r(x)$ *such that (21) is satisfied for every* k*, and*

$$\int_0^\infty f_r(x,t) = n_r(t), \qquad \forall t. \tag{26}$$

Proof: See the Appendix.

The PDE (25) is solved with the method of characteristics. We look, that is, for a function $\xi_r(t)$ such that along the *characteristic* trajectories $(x,t) = (\xi_r(t), t)$ the evolution of f_r is described by the simple ODE :

$$\frac{d}{dt} f_r(\xi_r(t), t) - f_r(\xi_r(t), t) = \frac{\partial}{\partial t} f_r - (x - 2\eta n_r) \frac{\partial}{\partial x} f_r - f_r = 0. \tag{27}$$

The solution thus obtained can then be used to solve for the degree distribution in each class r, under the assumption that the time scale τ in which the class sizes $n_r(\tau)$ change is much slower than the time scale t for network adjustment – i.e., specifically, we consider the limit $\tilde{\nu} \to 0$ with $\tau = \tilde{\nu} t$ finite. The derivations involved in the whole procedure are described in the Appendix, as part of the proof leading to the following result.

Proposition 8 *Let* $\tau = \tilde{\nu} t > 0$ *be given. Then, in the limit* $\tilde{\nu} \to 0$*, the degree distribution converges in each component to the Poisson distribution with mean* $2\eta n_r(\tau)$*, i.e.*

$$\lim_{\tilde{\nu} \to 0} n_{r,k}(t = \tau/\tilde{\nu}) = n_r(\tau) \frac{[2\eta n_r(\tau)]^k}{k!} e^{-2\eta n_r(\tau)} \tag{28}$$

The previous result indicates that the network topology eventually converges to that of a Poisson random network if agents exhibit large status-quo inertia and thus class sizes change slowly. The same applies (for every value of $\tilde{\nu}$) when the process approaches a stable configuration, since this also brings about a slow change in class sizes. In all of these cases, however, the convergence of the *degree* *distribution* is associated to an underlying network architecture in continuous flux that will typically be too intricate to be described other than statistically. It is in this sense that we can speak of a complex social network arising as the outcome of the evolutionary process.

4.4 The dynamics of class sizes

Now we turn to studying the evolution of class sizes, still under the simplifying assumption that their change proceeds within a time scale τ that is much longer than that within which the network adjusts. Then, we can rewrite the dynamics of the population sizes $n_r(\tau)$ as

$$\frac{dn_r}{d\tau} = \sum_{s=1}^{q} [n_{s,0} - n_{r,0}] \tag{29}$$

and, by virtue of the results of the previous subsection, use $n_{r,0} \cong n_r(\tau)e^{-2n_r(\tau)}$ as an accurate approximation for small $\tilde{\nu}$. This allows us to derive the following strong result on the long-run behavior of the process.

Proposition 9 *Let $\tau = \tilde{\nu}t$ be the relationship between the two time scales of the MFD governing action and network change. Then, in the limit $\tilde{\nu} \to 0$, the function $f(\mathbf{n})$ defined in Proposition 4 satisfies $df(\mathbf{n}(\tau))/d\tau \leq 0$, where it may hold with equality only if $\mathbf{n}(\tau)$ is a critical point of f (i.e. satisfies the FOC (12)).*

The above result indicates that the function $f(\cdot)$ plays a key role not only in the ultralong run analysis of the model but also in its long-run dynamics. For, on the one hand, recall from Section 3 that the global minima of the function f single out the ultralong-run configurations of the process. And, on the other hand, Proposition 9 establishes that it is a Lyapunov function for the MFD, the trajectories always moving towards lower values of f except when the system lies at a critical point of the function. The latter conclusion, in particular, leads to a further important consequence:

Corollary 10 *Under the conditions considered in Proposition 9, the local minima of the function f are asymptotically stable configurations of the MFD.*

The above corollary implies (recall our discussion in Subsection 4.1) that, if the population is large enough, the process can be predicted to spend an arbitrary long (but finite) time in any small neighborhood of a local minima of the function f with high probability. In this sense, we can view each of the local minima of f as an alternative long-run prediction if the process starts close to it. This provides a formal basis for our former heuristic discussion on the dichotomy between the long- and ultralong-run outcomes, as respectively embodied by *local* versus *global* minima of the function f.

To illustrate that dichotomy in a sharp way, it is useful to focus on how the alternative equilibria translate into different levels of overall connectivity. First, let us consider the *long-run* predictions of the model, as embodied by the asymptotically stable configurations of the MFD. At any such configuration $\mathbf{n}$ we can compute the average degree

$$z = \sum n_r z_r = 2\eta \left[L_+ n_+^2 + (q - L_+) n_-^2 \right] \tag{30}$$

where $z_r = 2\eta n_r$ is the average degree of nodes in component r, in the stationary state, and n_+, n_-, and L_+ (as functions of η) are precisely as in Proposition 5 and the associated discussion. Here we rely on the fact (that is straightforward to check directly) that the equilibria of the MFD yield aggregate configurations that have exactly the same structure as the minima of the function f. Thus, associated to the thresholds $\check{\eta}$ and $\hat{\eta}$ contemplated in Proposition 5, our analysis predicts three different regions in η where qualitatively very different long-run equilibria can arise, in turn inducing very different levels of network connectivity. For $q = 10$, the situation is graphically depicted in Figure 3.

The solid curve in the diagram traces the average degree induced by the long-run solutions derived from the mean-field model, for the whole range of η. First, we find that if $\eta < \hat{\eta}(q) = q/2$, there is an equilibrium, corresponding to $L_+ = 0$, where the network is an Erdös-Renyi random graph with an average degree that is relatively low, $z^* = 2\eta/q$, and slowly increasing with η. In particular, $z^* < 1$ which means that the network is fragmented in many small disconnected components – cf. Erdös and Rényi (1960), or Newman *et al.* (2001). On the other hand, if $\eta > \check{\eta}(q)$ ($\simeq 2.28$ for $q = 10$) there is an equilibrium with $L_+ = 1$ where the average degree z^* is relatively large and increases steeply as η grows. In particular, the fact that $z_r^* = 2\eta n_+ > 1$ for the large action class implies that a giant component exists that connects a significant fraction of the nodes in the network. Finally, we observe that since $\hat{\eta} > \check{\eta}$, there is a middle range $(\check{\eta}, \hat{\eta})$ ($= (2.28, 5)$ for $q = 10$) where the two kind of equilibria are possible:

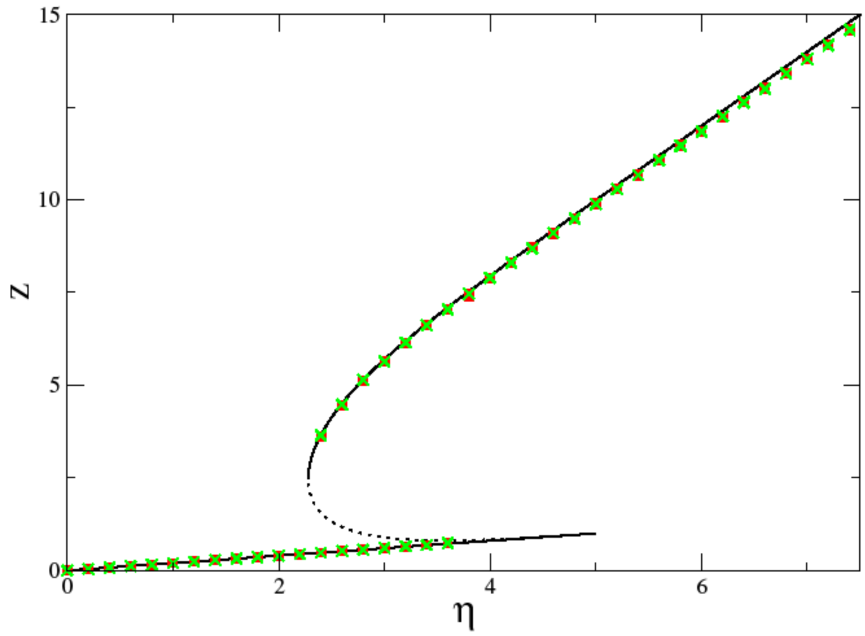

Figure 3: The solid curve represents the mean connectivity induced by the theoretical mean-field model (as given by (30)) against the rate of link formation, for a value of $q = 10$. The numerical simulations were conducted for two different values $\tilde{\nu}$ – i.e. $\tilde{\nu} = 1$ (squares) and $\tilde{\nu} = 10$ (crosses). The simulations are performed by implementing a series of gradual changes in η (starting from both high and low values for it), letting the system equilibrate at each stage before starting to record the situation. In the simulations, the low equilibrium becomes unstable below the theoretically predicted value of 5 because fluctuations have significant effects close to the transition point given the finiteness of the population. Such finiteness also explains the slight downward deviations observed along the upper branch for high values of η.

one with a low connectivity, fragmented network and wide action heterogeneity, another with much higher degree, a giant component, and significant action homogeneity.

Figure 3 also depicts the outcome of numerical simulations, which were conducted for a population size $N = 1000$, the indicated value of $q = 10$, and the two values of $\tilde{\nu} = 1$ and $\tilde{\nu} = 10$. We find that the asymptotically stable configurations singled out by the mean-field theory indeed act as strong local attractors of the simulations. It is interesting to observe that only a moderately large population size is enough to achieve a good predictive performance of the theoretical model. We also find that the predictions are essentially unaffected by the specific values of $\tilde{\nu}$ considered, even though they differ in an order of magnitude. The latter is in consonance with the minor role played by this parameter in the core of our analysis – e.g. in the characterization of the invariant distribution of the underlying stochastic process.

The long-run behavior of the system illustrated in Figure 3 displays several features worth highlighting. First, we observe that in the borders of the coexistence region marked by $\check{\eta}$ and $\hat{\eta}$, the configuration of the system may change discontinuously in response to small changes in η. Thus, for example, as this parameter increases slightly from just below $\hat{\eta} = q/2$ to a value above it, there is sharp increase in the average degree of the underlying network if the system was originally set at a low-connectivity configuration – i.e. one associated to the lower locus (or branch) of equilibria. Reciprocally, a similarly discontinuous transition also occurs at the lower border of the coexistence region as η falls below $\check{\eta}$, but now from a high- to a low-connectivity configuration. In fact, such "punctuations" in social behavior are found to arise in a number of different real-world contexts, where homophily and conformity are important reinforcing forces at work.[15] Although our model is too abstract and stylized to be directly usable for empirical analysis, it highlights a general mechanism of cross and cumulative feedback along those two dimensions (homophily and conformity)

[15] By way of illustration, we can focus on empirical evidence available on so-called knowledge networks, i.e. networks of collaboration based on the sharing of knowledge and expertise. Thus, for R&D partnerships among firms, Hagerdoorn (2002) reports a ten-fold increase during the decade 1980-1990. Or, concerning collaboration in academic research, Goyal *et al.* (2006) report a similarly steep (three-fold) increase in the per capita number of collaborations among academic economists in the last three decades. Finally, a similar sense of a rapidly unfolding process of network formation is gained from, say, the excellent account found in Castilla *et al.* (2000) of the rise to prominence of one of the most paradigmatic industrial districts of modern times: the semi-conductor firms located in Silicon Valley. They explain that, in roughly five years after the creation of Intel in 1968, much of the density of interpersonal connections that underlay the lively flow of ideas across the Valley was largely in place.

that may be at work in many such situations.

In addition, the model helps understand why many of those transitions in the real world often happen to lead into a resilient and rather persistent (or long-run) state of affairs. Relatedly, it also explains why history may matter. For, in our theoretical framework, once the discontinuous shift to a new configuration has taken place as a response to a change in our leading parameter η, even if this parameter were to gradually revert to its original value, the transition cannot be locally (i.e. "easily") overturned. The reason is that, along the new locus of equilibria thus reached, the corresponding equilibria are dynamically stable. Or, in the language of dynamical systems, the process can be said to display *hysteresis*. Conceptually, what this property reflects in our case is a qualitative change in structure – specifically, in the architecture of the social network – that makes the transition to the new configuration a robust outcome. As a consequence, it turns out that history matters and multiple, very diverse, configurations may persist as a mere reflection of different past experience.[16]

Let us now turn to the *ultralong run* time scale, where the prediction concerns the infinite-time limit for any (finite) population size. In that asymptotic perspective, we know that matters are crucially different. For, in view of the ergodicity of the process, there is a unique prediction that is independent of initial conditions. Outside the long-run coexistence region – i.e. if $\eta < 2.28$ or $\eta > 5$ – the issue is straightforward since even the long-run analysis provides a unique prediction. But, within that region – i.e. for $\eta \in (2.28, 5)$ – it all hinges upon how the prevailing rate of link formation η compares with the threshold $\eta^* = [(q-1)/(q-2)] \log(q-1) \simeq 2.47$. Below this threshold, the local minimum of f where $L_+ = 0$ is also the global minimum (cf. (14)) and therefore it gives rise to the unique ultralong-run configuration. In this case, therefore, the system displays minimum conformity and a relatively low connectivity in the ultralong run – i.e. the outcome is located along the lower branch in Figure 3. The opposite occurs when $\eta > \eta^*$, in which case the ultralong run outcome lies on the upper branch with $L_+ = 1$, so that there is relatively strong conformity

[16] As a follow-up on Footnote 15 – where we briefly discussed sharp transitions in knowledge networks – let us now complement that discussion with some informal reference to related evidence that suggests equilibrium multiplicity. For example, concerning R&D interfirm networks, Hagerdoorn (2002) explains that similar sectors in the same countries, or the same sectors in different countries, often display drastically different levels of cooperation (i.e. of connectivity) in their R&D partnership network. Or, concerning the rise of industrial districts, one may refer to the celebrated account by Saxenian (1994) that compares the developments of Silicon Valley and Massachusetts' Route 128. She highlights that, despite the fact that both cases involved firms operating in the same industry (that were also direct competitors), the networks arising in each region were markedly different.

and a high average degree.

5 Summary and Conclusions

The main insight delivered by our model can be succinctly described as follows. In a scenario where the social network (co-)evolves alongside agents' efforts to coordinate with their partners, the interplay between these two dimensions can generate rich dynamics in both the social network and agents' behavior. The focus of the model has not been on the coordination issue *per se* but on how agents' efforts to align their game behavior affects the long-run network topology. We have seen, specifically, that small changes in the underlying parameters can bring about sharp and robust transitions between a sparse network and a connected one with complex architecture, in either direction. This, in turn, is associated to analogous shifts in the extent of social conformity/homogeneity – which is induced, and also itself induces, the change in network structure. Finally, we have shown that the nature of such transitions, and their interpretation, is crucially affected by the time horizon under consideration. In the long run (for large finite times), history matters and multiplicity can arise. Instead, in the ultralong run (for infinite times), the ergodicity of the process leads to unique outcomes independently of initial conditions.

The model is simple and stylized but its behavior matches some of the features that have been highlighted as important for a number of social contexts. The objective of the paper, however, is inherently theoretical and thus we have not attempted to address empirical issues in any detail. Rather, our emphasis has been to develop a full-fledged analysis of the complex dynamics of the model, showing that the two simple forces of homophily and conformity may generate rich and interesting social dynamics. A second objective of the paper has been of a methodological nature. Namely, we have illustrated that some mathematical tools (e.g. ergodic and mean-field analysis) that have been widely used to study contexts with simple or unstructured matching patterns for the population can be significantly extended. Specifically, they can be applied to setups where the pattern of interaction itself (i.e. the network) displays a complex architecture and coevolves with agents' behavior.

Appendix

Proof of Proposition 1: First, we argue that, for any two states $\omega, \omega' \in \hat{\Omega}$, it is possible to find a finite sequence of transitions (jointly occurring with some positive probability, bounded away from zero) that leads from one to the other, in a finite time. To see this, note that there is such positive probability for a transition to occur from ω to a state $\tilde{\omega}$ consisting of the empty network (i.e. with no links) and an action profile where every agents chooses the same action as in ω'. Indeed, such a transition takes place if all links in ω vanish and then all nodes in ω with an action different from the one they display in ω', receive an action-revision opportunity and switch to the latter. If $\omega' \in \hat{\Omega}$, there must also be a transition from $\tilde{\omega}$ to ω' occurring with positive probability. Indeed all links present in ω' can be added sequentially, starting from the state $\tilde{\omega}$. Indeed, by construction, these links are all between agents displaying the same action. This implies that a transition from ω to ω' occurs with positive probability, as claimed.

The previous argument establishes that all states in $\hat{\Omega}$ belong to a single recurrent class and, therefore, the Markov process has a unique invariant distribution μ. Next we show that all states $\omega \notin \hat{\Omega}$ are transient. To this end, we first argue that the probability of a transition from a state $\omega \in \hat{\Omega}$ to any $\omega' \notin \hat{\Omega}$ is zero. This readily follows from the following two facts, which hold at every t: (a) from the mechanism of *link formation*, only agents currently choosing the same action can create a new link at t; (b) from the mechanism of *action revision*, no node belonging to a network component whose agents all currently choose the same action will change her action at t. More specifically, take any state $\omega \in \hat{\Omega}$ and consider the possible transitions away from it that pertain to some nodes i and j. We need to consider three possibilities: *i)* $g_{ij} = 0$ and $\alpha_i \neq \alpha_j$, in which case the link ij will not form, *ii)* the link ij is present ($g_{ij} = 1$) and $\alpha_i = \alpha_j$, in which case agents will not change actions, or *iii)* $g_{ij} = 0$ and $\alpha_i = \alpha_j$, in which case either the link ij forms or agents revise actions, so that one is back to one of the previous two possibilities. In none of the cases, a possible transition can bring the process to a state not in $\hat{\Omega}$. Thus, since any state $\tilde{\omega}$ with an empty network belongs to $\hat{\Omega}$, the argument in the first part of the proof implies that every state $\omega' \notin \hat{\Omega}$ is transient and therefore $\mu(\hat{\Omega}) = 1$. ■

Proof of Proposition 2: Clearly, a sufficient condition for the stationarity embodied by (4) is given by the following "detailed-balance" equalities [19]:

$$\mu(\omega)\ \rho(\omega \to \omega') = \mu(\omega')\ \rho(\omega' \to \omega) \qquad (\omega, \omega' \in \Omega). \tag{31}$$

Thus let us verify that the above equalities hold for the probability distribution μ given by (4). First we notice that, for any pair of states ω', ω, either $\rho(\omega' \to \omega)$ and $\rho(\omega \to \omega')$ are both zero, or they are both non-zero. Hence we only need to check (31) for states ω, ω' across which a direct transition is possible with positive probability, i.e. $\rho(\omega \to \omega') > 0$. These possible transitions must fall into two categories: link adjustment alone (with actions remaining fixed), or action adjustment alone (with links unchanged). We address each of them in turn.

For the first case (link adjustment), we need to consider any two states, $\omega = (\boldsymbol{\alpha}, G)$ and $\omega' = (\boldsymbol{\alpha}', G')$, such that $\boldsymbol{\alpha} = \boldsymbol{\alpha}'$ and $g_{ij} = g'_{ij}$ for all $i, j \in P$ except for one pair, k and ℓ, such that $\alpha_k = \alpha_\ell$ and, say, $g_{k\ell} = 1$ but $g'_{k\ell} = 0$. Then, from (4), we have:

$$\frac{\mu(\omega)}{\mu(\omega')} = \frac{2\eta}{N-1}.$$

On the other hand, note that the rate $\rho(\omega' \to \omega) = 2\eta/(N-1)$ since a link-creation opportunity arrives to either k or ℓ at the rate 2η and the link $k\ell$ is actually created if the one who receives the opportunity meets the other one – an event with probability $1/(N-1)$. On the other hand, the rate $\rho(\omega' \to \omega) = 1$, since each existing link vanishes at the rate $\lambda = 1$. Combining these consideration, we arrive at the following conclusion:

$$\frac{\mu(\omega)}{\mu(\omega')} = \frac{2\eta}{N-1} = \frac{\rho(\omega' \to \omega)}{\rho(\omega \to \omega')},$$

which obviously implies the detailed-balance condition (31).

For action revision, we need to consider any two states, $\omega = (\boldsymbol{\alpha}, G)$ and $\omega' = (\boldsymbol{\alpha}', G')$, such that $G = G'$ and $\alpha_i = \alpha'_i$ for all $i \neq k$, for some $k \in P$, whereas $\alpha_k \neq \alpha'_k$. The rate of this transition is non-zero only if agent k has no links ($g_{ik} = 0\ \forall i \in P/\{k\}$) and, in this case, $\rho(\omega \to \omega') = \rho(\omega' \to \omega) = \tilde{\nu} \equiv (1-u)\nu$. Clearly $\mu(\omega)$ of (4) satisfies (31) for pairs of states of this type.

This implies that the probability distribution μ given by (4) is invariant, and indeed the unique one in view of ergodicity. ■

Proof of Proposition 4: We rely on repeated applications of Stirling's formula $k! \simeq (k/e)^k$, for large k:

$$\begin{aligned}\log \mu(\hat{\Omega}(\mathbf{N})) &\simeq \log \Upsilon - \sum_{r=1}^{q} \left[N_r \log(N_r/N) - \frac{N_r(N_r-1)}{2} \log\left(1+\frac{2\eta}{N-1}\right)\right] \\ &\simeq \log \Upsilon - N\left[f(\mathbf{N}/N) + O(1/N)\right]. \end{aligned} \tag{32}$$

where f is defined in (9). The proof is completed by showing that $N^{-1}\log\Upsilon \to f_0$, for some finite constant f_0. This is done by using the normalization identity

$$-\log \Upsilon = \log\left\{\sum_{\mathbf{N}} \frac{N!}{\prod_{r=1}^{q} N_r!} \prod_{r=1}^{q}\left[\left(1+\frac{2\eta}{N-1}\right)^{\frac{1}{2}N_r(N_r-1)}\right]\right\} \tag{33}$$

where the sum runs on all sets of positive integers $\mathbf{N} = (N_1, \ldots, N_q)$ such that $N_1 + N_2 + \ldots + N_q = N$. We easily derive a lower bound $-\log\Upsilon \geq N\log q$, by setting $\eta = 0$ and observing that the sum becomes the multinomial expansion of q^N. For the upper bound, we use $N_r(N_r-1) \leq N(N-1)$ in the exponent of (33) so that $-\log\Upsilon \leq \frac{qN(N-1)}{2}\log\left(1+2\eta/(N-1)\right) + N\log q$. Using $\log(1+x) \leq x$ we finally arrive at

$$\log q \leq -\frac{1}{N}\log\Upsilon \leq q\eta + \log q$$

which implies that f_0 is finite, as claimed. ■

Proof of Proposition 5: Let $\mathbf{n}^*$ be a configuration that satisfies the FOC (12). In order for it to be a minimum of f, we need to check that f increases along all directions, on the simplex Δ^{q-1} around $\mathbf{n}^*$ (second-order conditions). For an infinitesimal perturbations $\mathbf{n} = \mathbf{n}^* + \boldsymbol{\varepsilon} \in \Delta^{q-1}$, the vector $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_q) \in \mathbb{R}^q$ satisfies $\sum_{r=1}^{q} \varepsilon_r \equiv 0$, and the change in the value of f to leading order is:

$$f(\mathbf{n}) \simeq f(\mathbf{n}^*) + \frac{1}{2}\sum_{r=1}^{q} \frac{1-2\eta n_r^*}{n_r^*}\varepsilon_r^2 + O(\varepsilon^3) \tag{34}$$

A sufficient condition for $\mathbf{n}^*$ to be a minimum is that $f(\mathbf{n}) - f(\mathbf{n}^*) \geq 0$ for all possible vectors $\boldsymbol{\varepsilon}$ which satisfy $\sum_{r=1}^{q} \varepsilon_r \equiv 0$. Notice that, in view of the FOC, $2\eta n_- \leq 1 \leq 2\eta n_+$.

Let us divide our discussion into three different cases: $L_+ = 0$, $L_+ = 1$, $L_+ > 1$. We address each of them in turn.

In the case $L_+ = 0$, by symmetry, we must have $n_r^* = n_- = 1/q$ for all $r =$

$1, ..., q$. But, this is compatible with the FOC (12) if, and only if, $1/q \leq 1/(2\eta)$, or $\eta \leq q/2$. And in this range, the second term in (34) is clearly positive, so the second order conditions are also satisfied.

Next, we consider the case with $L_+ = 1$. Let $n_1 = n_+$ (i.e. the larger class adopts action a_1) and $n_r = n_- < n_+$ for $r > 1$. In all directions with $\epsilon_1 = 0$ (and $\sum_{r>1} \varepsilon_r = 0$), $f(\mathbf{n}^* + \boldsymbol{\varepsilon}) - f(\mathbf{n}^*) > 0$ because $2\eta n_- < 1$, as in the case with $L_+ = 0$. Therefore, we need consider only the direction $\varepsilon_1 \neq 0$ and $\varepsilon_r = -\varepsilon_1/(q-1)$ for $r > 1$ which is orthogonal to all vectors $\boldsymbol{\varepsilon} = (0, \varepsilon_2, \ldots, \varepsilon_q)$, with $\sum_{r>1} \varepsilon_r = 0$. For these, the variation of f can be seen to take the form

$$\begin{aligned} \frac{1}{2}\sum_{r=1}^{q} \frac{1-2\eta n_r^*}{n_r^*}\varepsilon_r^2 &= \left[\frac{1}{n_+(1-n_+)} - \frac{2\eta q}{q-1}\right]\varepsilon_1^2. \\ &= 2\frac{qn_+ - 1}{q-1}\left(\frac{dn_+}{d\eta}\right)^{-1}\varepsilon_1^2. \end{aligned} \tag{35}$$

The last expression is readily confirmed by taking the derivative of (13) with respect to η. This yields

$$\frac{dn_+}{d\eta} = 2\frac{n_+(1-n_+)}{q-1}\left[qn_+ - 1 + \eta q\frac{dn_+}{d\eta}\right]$$

which is easily solved for $\frac{dn_+}{d\eta}$. Then, (35) completes the proof of the case $L_+ = 1$, since it shows that the solution is a minimum if and only if n_+ increases with η.

Finally, in order to prove that solutions of the FOC with $L_+ > 1$ are not minima, it is sufficient to find a direction $\boldsymbol{\varepsilon}$ along which f decreases. Let $n_r = n_+$ for $r \leq L_+$ and $n_r = n_-$ otherwise. Along the direction $\boldsymbol{\varepsilon} = (\epsilon_1, -\epsilon_1, 0, \ldots, 0)$ the variation in f is given by

$$f(\mathbf{n}) - f(\mathbf{n}^*) = -\frac{2\eta n_+ - 1}{n_+}\epsilon_1^2 + O(\boldsymbol{\varepsilon}^3)$$

This is negative because $2\eta n_+ > 1$, which concludes the proof. ■

Proof of Proposition 6: Let $N_{r,k} \equiv n_{r,k}N$ stand for the *total number* of nodes displaying each action r and degree k. Consider first the case where $k > 0$ and assume, for simplicity, that the process has already abandoned transient states and all links connect agents choosing the same action. Then, over time, the magnitudes $N_{r,k}$ change solely due to link creation and link destruction. In any time interval $[t, t+\Delta t]$ of infinitesimal length Δt, link creation opportunities arrive independently to each node with probability $\eta\Delta t$, while each existing link is destroyed with probability $\lambda\Delta t = \Delta t$ (recall the normalization $\lambda = 1$). Then, we now argue that expected change $E[\Delta N_{r,k}]$ over that infinitesimal time interval is given by the following expression:

$$E[\Delta N_{r,k}] = N\eta \left\{ 2n_{r,k-1} \sum\nolimits_{k'} n_{r,k'} - 2n_{r,k} \sum\nolimits_{k'} n_{r,k'} \right\} \Delta t \qquad (36)$$
$$+ N \left\{ (k+1)n_{r,k+1} - kn_{r,k} \right\} \Delta t \qquad (k > 0).$$

The first bracketed term of (36) concerns events of link creation. These events affect $N_{r,k}$ through five possible routes:

(i) Some node counted in $N_{r,k-1}$ is selected for link creation and then meets another node counted in $N_{r,k-1}$ as well. This occurs with probability $\eta\Delta t\ N\ n_{r,k-1}\left[(N_{r,k-1}-1)/N\right] = N\eta\left[n_{r,k-1}^2 + \mathcal{O}(1/N)\right]\Delta t$.

(ii) Some node counted in $N_{r,k-1}$ is selected for link creation and then meets another node counted in $N_{r,k'}$ for k' such that $k \neq k' \neq k-1$, or *vice versa*. This occurs with probability $2\eta\Delta t\ N\ n_{r,k-1}\left[(\sum_{k\neq k'\neq k-1} N_{r,k'})/N\right] = 2N\eta\left[n_{r,k-1}(\sum_{k\neq k'\neq k-1} n_{r,k'})\right]\Delta t$.

(iii) Some node counted in $N_{r,k}$ is selected for link creation and then meets another node counted in $N_{r,k}$ as well. This occurs with probability
$\eta\Delta t\ N\ n_{r,k}\left[(N_{r,k}-1)/N\right] = N\eta\left[n_{r,k}^2 + \mathcal{O}(1/N)\right]\Delta t$.

(iv) Some node counted in $N_{r,k}$ is selected for link creation and then meets another node counted in $N_{r,k'}$ for k' such that $k \neq k' \neq k-1$, or *vice versa.* This occurs with probability $2\eta\Delta t\ N\ n_{r,k}\left[(\sum_{k\neq k'\neq k-1} N_{r,k'})/N\right] =$
$2N\eta\left[n_{r,k}(\sum_{k\neq k'\neq k-1} n_{r,k'})\right]\Delta t$.

(v) Some node counted in $N_{r,k}$ is selected for link creation and then meets another node counted in $N_{r,k-1}$, or *vice versa.* This occurs with probability $2\eta\Delta t\ N\ n_{r,k}\left[N_{r,k-1}/N\right] = 2N\eta\left[n_{r,k}n_{r,k-1}\right]\Delta t$.

Now let us determine what is the induced change in $N_{r,k}$ for each of the above possibilities. For (i), $\Delta N_{r,k} = 2$ since the link created brings in two new nodes to the set of those that display action a_r and have degree k; for (ii), $\Delta N_{r,k} = 1$ since only one new node is added to that set; for (iii), $\Delta N_{r,k} = -2$ since the link created has the two connecting nodes increase their degree to $k+1$ and thus abandon the set in question; for (iv), $\Delta N_{r,k} = -1$ since the the connecting node that originally had degree k then has degree $k+1$ and thus abandons the set; for (v), $\Delta N_{r,k} = 0$ since the entry of one node in the set is exactly cancelled by the exit of one other node. Bringing together all these considerations, the first bracketed term of (36) readily obtains.

Let us now address the second bracketed term of (36). In this respect, simply note that a single link of some node with k' links is removed with a probability $\Delta t\ k' N_{r,k'} = N\ k' n_{r,k'}\ \Delta t$. Then, the desired expression simply follows from the observation that the removal of one link from a node with degree $k+1$ induces $\Delta N_{r,k} = 1$ (i.e. increases $N_{r,k}$ by one), while if it affects a node with degree k it leads to $\Delta N_{r,k} = -1$.

Next, we consider the case where $k = 0$ and compute the expected change in the numbers $N_{r,0}$ for each r. Unlike for the case with $k > 0$, the dynamics is now affected by action adjustment, thus giving rise to the following expression:

$$E[\Delta N_{r,0}] = N\eta \left\{ -2n_{r,0} \sum\nolimits_{k'} n_{r,k'} \right\} \Delta t + N n_{r,1} \Delta t + N(1-u)\nu \sum_{s=1}^{q} [n_{s,0} - n_{r,0}]\, \Delta t. \tag{37}$$

The first two terms in (37) are just as before – reflecting link creation and link destruction, respectively – except that they now can operate only in one direction: neither isolate nodes can lose any links, nor link creation can lead a node to become an isolate node. The third term, on the other hand, embodies the process unfolding at the rate $(1-u)\nu$ that makes isolate nodes drift across the set of possible actions.

Expressions (36)-(37) give the expected change in the absolute numbers of nodes $N_{r,k}$ displaying each action r and degree k. The corresponding change in frequencies $n_{r,k}$ obviously satisfies $E[\Delta N_{r,k}] = N\ E[\Delta n_{r,k}]$, thus yielding (16)-(17), as desired. ■

Proof of Proposition 7: For each r and $k > 0$, the representation of $F_{r,k}(\mathbf{n})$ in terms of $f_r(x,t)$ reads:

$$\begin{aligned} F_{r,k}(\mathbf{n}) &= \int_0^\infty dx f_r(x,t)(x - 2\eta n_r) \left[\frac{x^k}{k!} - \frac{x^{k-1}}{(k-1)!} \right] e^{-x} \\ &= -\int_0^\infty dx f_r(x,t)(x - 2\eta n_r) \frac{\partial}{\partial x} \left[\frac{x^k}{k!} e^{-x} \right] \\ &= \int_0^\infty dx \frac{x^k}{k!} e^{-x} \frac{\partial}{\partial x} \left[f_r(x,t)(x - 2\eta n_r) \right] \end{aligned}$$

where, in the last passage, we have integrated by parts using the fact that

$f_r(x,t)(x-2\eta n_r)\frac{x^k}{k!}e^{-x}$ vanishes both at $x=0$, for all $k>0$, and at $x\to\infty$. Therefore the ODE $\dot n_{r,k} = F_{r,k}(\mathbf{n})$ implies that for all r and $k>0$

$$0 = \int_0^\infty dx \frac{x^k}{k!} e^{-x} \left[\partial_t f_r - \partial_x \left(x - 2\eta n_r\right) f_r\right]. \tag{38}$$

If (38) holds for all $k>0$, then it means that the term in square brackets vanishes, i.e. f_r satisfies (25). Since the transformation to the Poisson representation is invertible, for any initial conditions specifying the $n_{r,k}(0)$ one can identify a suitable initial condition $\phi_r(x)$ for each f. ■

Proof of Proposition 8: First, we determine the characteristic paths $(\xi_r(t),t)$ along which the PDE (25) is equivalent to the ODE (27). These characteristic trajectories are obtained from a solution to the equation

$$\frac{d\xi_r(t)}{dt} = 2\eta n_r(t) - \xi_r(t), \quad \xi_r(0) = x_0,$$

i.e.

$$\xi_r(t) = x_0 e^{-t} + 2\eta \int_0^t ds e^{s-t} n_r(s) \equiv e^{-t}x_0 + \chi_r(t). \tag{39}$$

where the last equality defines

$$\chi_r(t) \equiv 2\eta \int_0^t ds n_r(s) e^{s-t}. \tag{40}$$

On the characteristic path starting from the initial condition $(x_0, t=0)$, the function f_r satisfies $\frac{df_r}{dt} = f_r$, i.e. $f_r(\xi_r(t),t) = \phi_r(x_0)e^t$. Inverting (39) we obtain $x_0 = e^t[x - \chi_r(t)]$ and therefore:

$$f_r(x,t) = e^t \phi_r \left\{ e^t [x - \chi_r(t)] \right\}. \tag{41}$$

For $x \ge \chi_r(t)$, this solution is related to values of the initial condition $\phi_r(x_0)$ with $x_0 \ge 0$. Instead, for $x < \chi_r(t)$, the initial condition is associated to negative values of x_0, for which ϕ_r is not defined. We therefore set the values of $\phi_r(x_0)$ for $x_0 < 0$ in such a way as to satisfy the condition (26). In order to do this, integrate (41) over $x \in [0, \infty)$:

$$n_r(t) = \int_0^\infty dx e^t \phi_r \left\{ e^t [x - \chi_r(t)] \right\} = \int_{-e^t \chi_r(t)}^\infty dz \phi_r(z) \tag{42}$$

$$= \int_{-e^t \chi_r(t)}^0 dz \phi_r(z) + n_r(0) \tag{43}$$

where we have changed variable to $z = e^t [x - \chi_r(t)]$ and split the integrals for $z \geq 0$ and $z < 0$. Taking now the derivative with respect to t on both sides, we get

$$\frac{dn_r}{dt} = \phi_r(-e^t \chi_r) \frac{d}{dt} \left[e^t \chi_r(t) \right] = 2\eta \phi(-e^t \chi_r) e^t n_r(t) \tag{44}$$

or

$$\phi(-e^t \bar{\chi}) = \frac{e^{-t}}{2\eta} \frac{d}{dt} \log n_r(t).$$

This implicitly defines the function $\phi_r(x_0)$ for negative x_0 in terms of the function $n_r(t)$. Notice that $e^t \chi_r(t)$ is an increasing function of t as long as $n_r > 0$, so the procedure yields a unique selection.

The solution of the fraction $n_{r,k}$ of nodes in class r with degree k can then be expressed, with the same change of variable as before, as

$$n_{r,k}(t) = \int_{-e^t \chi_r}^\infty dz \phi_r(z) \left\{ \frac{[\chi_r + (z - \chi_r) e^{-t}]^k}{k!} e^{-\chi_r + (\chi_r - z) e^{-t}} \right\} \tag{45}$$

where $\chi_r(t)$ is a function of $n_r(t)$.

Now, we specialize this solution to the case where the evolution of n_r takes place on time scales much longer that that of $n_{r,k}$. To this end, let the population in component r be a smooth (continuous, differentiable) function $n_r(\tau)$ of the time variable $\tau = \tilde{\nu} t$. Then, by (40), for $t = \tau / \nu$ and small ν:

$$\chi_r(t) = 2\eta n_r(\tau) - 2\eta \tilde{\nu} \frac{dn_r}{d\tau} + O(\tilde{\nu}^2).$$

Likewise the term in braces of (45) converges uniformly to $\frac{(2\eta n_r)^k}{k!} e^{-2\eta n_r}$ in the limit $\tilde{\nu} \to 0$ with $\tau = \tilde{\nu} t > 0$ given. Since, by (42), the remaining integral on z is exactly $n_r(\tau)$, the desired conclusion follows. ■

Proof of Proposition 9: The derivative of f with respect to τ is given by

$$\frac{df}{d\tau} = \sum_r [1 + \log n_r - 2\eta n_r] \frac{dn_r}{d\tau} \tag{46}$$

$$= \sum_r [1 + \log n_{r,0}] \sum_s [n_{s,0} - n_{r,0}] \tag{47}$$

$$= -\frac{1}{2} \sum_{r,s} [\log n_{s,0} - \log n_{r,0}] [n_{s,0} - n_{r,0}] \leq 0 \tag{48}$$

where we use (29) and also $n_{r,0} = n_r e^{-2\eta n_r}$, the latter obtained from Proposition 8 by particularizing (28) to $k = 0$. Finally, in the last line we split the sum $\sum_{r,s}[\ldots] = \frac{1}{2}\sum_{r,s}[\ldots] + \frac{1}{2}\sum_{r,s}[\ldots]$, we interchanged the indices in the second term, and recombined the resulting expression. Given that $\log x$ is an increasing function, the expression (48) is non-negative, and it is zero only if $n_{r,0} = n_{s,0}$ for all $r, s = 1, \ldots, q$. This establishes the required property for time derivative of f in the slow time scale τ of the MFD. ■

References

[1] van Alstyne, M. (1997): "The state of network organization: a survey in three frameworks," *Journal of Organizational Computing and Electronic Commerce* 7, 83-151.

[2] Bala, V. and S. Goyal (1999): "A non-cooperative model of network formation," *Econometrica* **68**, 1181-1230.

[3] Barabási, A.-L. and R. Albert (1999):"Emergence of scaling in random networks", *Science* **286**, 509-12.

[4] Benaïm, M. and J. Weibull (2003*a*): "Deterministic approximation of stochastic evolution in games," *Econometrica* **3**, 873-903.

[5] Benaïm, M. and J. Weibull (2003*b*): "Deterministic approximation of stochastic evolution in games – a generalization," *mimeo*, Université de Neuchatel and Boston University.

[6] Binmore, K., L. Samuelson, and R. Vaughan (1995): "Musical Chairs: Modelling Noisy Evolution," *Games and Economic Behavior* **11**, 1-35.

[7] Blume, L. (1993): "The statistical mechanics of strategic interaction", *Games and Economic Behavior* **4**, 387-424.

[8] Bollobás, B. (2001): *Random Graphs* (2nd. Edition; First Edition 1985), Cambridge (U.K.): Cambridge University Press.

[9] Burt, R. S. (1992): *Structural Holes: The Social Structure of Competition*, Cambridge: Harvard University Press.

[10] Calvó-Armengol, A. and M. O. Jackson (2004): "The effects of social networks on employment and inequality," *American Economic Review* **94**, 426-454.

[11] Castells, M. (1996): *The Information Age: Economy, Society, and Culture, Volume I: The Rise of the Network Society*, Massachusetts: Blackwell Publishers.

[12] E. J. Castilla, H. Hwang, E. Granovetter, and M. Granovetter (2000): "Social Networks in Silicon Valley," in C.-M. Lee, W. F. Miller, M. G. Hancock, and H. S. Rowen, editors, *The Silicon Valley Edge*, Stanford: Stanford University Press.

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

[14] Ely, J. (2003): "Local Conventions," *Berkeley Electronic Press Journals, Advances in Theoretical Economics* **2**,1,1.

[15] Erdös, P. and A. Rényi (1959): "On random graphs I," *Publicationes Mathematicae Debrecen* **6**, 290-297.

[16] Erdös, P. and A. Rényi (1960): "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17-61.

[17] Ehrhardt, G. C. M. A. and M. Marsili (2005): "Potts model on random trees," *J. Stat. Mech.* P02006.

[18] Fafchamps, M. and S. Lund (2003): "Risk-sharing networks in rural Philippines," *Journal of Development Economics* **71**, 261-87.

[19] C.W. Gardiner (2004): *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Berlin: Springer-Verlag.

[20] Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, A. Shleifer (1992): "Growth in Cities," *The Journal of Political Economy* **100**, 1126-1152.

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

[22] Goyal S., M. van der Leijy, J. L. Moraga-González (2006): "Economics: an emerging small world?," *Journal of Political Economy* **114**, 403-412.

[23] Granovetter, M. (1974): *Getting a Job: a Study on Contacts and Careers*, Chicago: Chicago University Press.

[24] Granovetter (1985): "Economic action and social structure: the problem of embeddedness," *American Journal of Sociology* **91**, 481-510.

[25] Grossman, J. (2002): "The evolution of the mathematical research collaboration graph," *Congressus Numerantium* **158**, 201-212.

[26] Hagedoorn, J. (2002): "Inter-firm R&D partnerships: an overview of major trends and patterns since 1960", *Research Policy* **31**, 477-492.

[27] Jackson, M. O. (2005): "A survey of models of network formation: stability and efficiency," in Demange G. and M. Wooders (eds.), *Group Formation in Economics; Networks, Clubs and Coalitions*, Cambridge: Cambridge University Press.

[28] Jackson, M. O. and A. Wolinsky (1996): "A Strategic model of economic and social networks," *Journal of Economic Theory* **71**, 44-74.

[29] Jackson, M. O. and B. Rogers (2007): "Meeting Strangers and Friends of Friends: How Random are Socially Generated Networks?" *American Economic Review* **97**, 890-915.

[30] Jacobs, J. (1984): *Cities and the Wealth of Nations: Principles of Economic Life*, New York : Random House.

[31] Kirman, A., D. K. Herreiner, and G. Weisbuch (2000): "Market organization and trading relationships," *Economic Journal* **110**, 411-436.

[32] Kogut, B., P. Urso, and G. Walker (2006): "The emergent properties of a new financial market: American venture capital syndication from 1960 to 2005," forthcoming in *Management Science.*

[33] Krackhardt, D., and J. R. Hanson (1993): "Informal networks: the company behind the chart," *Harvard Business Review* **71**, 104-11.

[34] Kranton, R. and D. Minehart (2001): "A theory of buyer-seller networks," *American Economic Review* **91**, 485-508.

[35] Marsili, M., F. Vega-Redondo, and F. Slanina (2004): "The rise and fall of a networked society: a formal model," *Proceedings of the National Academy of Sciences, USA* **101**, 1439-43.

[36] McKelvey, R. D. and T. R. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and Economic Behavior* **10**, 6-38.

[37] Molloy M. and B. Reed (1995):"A critical point for random graphs with a given degree sequence", *Random Structures and Algorithms* **6**, 161-79.

[38] Montgomery, J. (1991): "Social networks and labor market outcomes: toward an economic analysis," *American Economic Review* **81**, 1408-1418.

[39] Murgai, R., P. Winters, E. Sadoulet, and A. de Janvry (2002): "Localized and incomplete mutual insurance," *Journal of Development Economics* **67**, 245-274.

[40] Newman, M. E. J. (2001): "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences USA* **98**, 404-409.

[41] Newman, M. E. J. (2003):"The structure and function of complex networks", *SIAM Review* **45**, 167-256.

[42] Newman, M.E., S.H. Strogatz, and D.J. Watts (2001): "Random graphs with arbitrary degree distributions and their applications," *Physical Review E* **64**, 026118.

[43] Powell, W. W., D. R. White, K. W. Koput, and J. Owen-Smith (2005): "Network dynamics and field evolution: the growth of inter-organizational collaboration in the life sciences," *American Journal of Sociology* **110**, 1132-1205.

[44] Saxenian, A. (1994): *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Cambridge, Mass.: Harvard University Press.

[45] Vega-Redondo, F. (2006): *Complex Social Networks*, forthcoming in the *Econometric Society Monograph Series*, Cambridge University Press.

[46] Watts D.J. and S. H. Strogatz (1998): "Collective dynamics of 'small-world' networks", *Nature* **393**, 440-42**.**

[47] Young, P. (1998): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton, NJ: Princeton University Press.