UNIVERSITY OF MANNHEIM

University of Mannheim / Department of Economics

Working Paper Series

Social networks and the process of "globalization"

Georg Duernecker Fernando Vega-Redondo

Working Paper 12-18

October 2012

Social networks and the process of "globalization"

Georg Duernecker
University of Mannheim

Fernando Vega-Redondo
European University Institute

October 8, 2012

Abstract

We propose a stylised dynamic model to understand the role of social networks in the phenomenon we call "globalization." This term refers to the process by which even agents who are geographically far apart come to interact, thus being able to overcome what would otherwise be a fast saturation of local opportunities. A key feature of our model is that the social network is the main channel through which agents exploit new opportunities. Therefore, only if the social network becomes global (heuristically, it "reaches far in few steps") can global interaction be steadily sustained. To shed light on the conditions under which such a transformation may, or may not, take place is the main objective of the paper.

One of the main insights arising from the model is that, in order for the social network to turn global, the economy needs to display a degree of "geographical cohesion" that is neither too high (for then global opportunities simply do not arise) nor too low (then the meeting mechanism displays too little structure for the process to take off). But if globalization does materialize, we show that it is a robust state of affairs that often arises abruptly as key parameters change. This occurs, in particular, as the rate of arrival of ideas rises, or when there is a high enough increase in the range at which the network transmits information.

Keywords: Social networks, Globalization, Search, Cooperation, Social Cohesion, Innovation.

JEL classif. codes: D83, D85, C73, O17, O43.

1 Introduction

The idea that most economies are fast becoming more globalized has become a commonplace, a *mantra* repeated by the press and popular literature alike as one of the distinct characteristics of our times. Economists, on the other hand, have also started to devote substantial effort to constructing measures of globalization that extend well beyond the traditional concern with trade openness (see Section 2 for a short summary of this predominantly empirical literature). Only relatively scant attention, however, has been devoted to understanding the phenomenon from a theoretical viewpoint. And this, of course, not only limits our grasp of matters at a conceptual level. It also limits our ability to advance further along the empirical route.

The main objective of this paper is to provide a terse model of globalization from the perspective afforded by the theory of social networks. A first question that arises concerns the nature of the phenomenon itself:

what is globalization? Our reply is that it involves two key (related) features. First, agents are relatively close in the social network, even if the population is very large – this is the situation often described as a "small world." Second, the links (connections) among agents tend to span long geographical distances.

A second question is equally basic: *Why is globalization important?* In response to this question, our model highlights the following points. Economic opportunities for collaboration appear both locally and globally (in terms of geographical space). The former are relatively easy to find and exploit but they are also limited in number. Hence sustained economic expansion must be able to access opportunities at a global scale. The social network is crucial in this respect, since information and trust (both of which underlie collaboration) are largely channelled through it. Sustained economic expansion, therefore, can only unfold if the social network itself becomes global in the sense suggested before. This, in sum, is the process of socio-economic globalization that provides the fuel and support for growth.

The former considerations raise a host of interesting issues. How and when is an economy able to build a truly global social network? Is the corresponding process gradual or abrupt? Does it lead to a robust state of affairs, or is it a fragile outcome hinging upon some quite special circumstances? What is the role of geography? Does geography provide a structure that facilitates, or instead hinders, globalization? Can we think of the process as some accumulation of social capital supporting economic activity? Is the there a role for policy? Are there temporary policy measures that may achieve permanent results?

In attempting to address the previous questions, our model provides some novel insights into the phenomenon. We end this short introduction by summarizing some of them. In general, if the economic potential associated to globalization is large, its materialization arises abruptly (i.e. as a sharp response to small changes in the environment). The transition, moreover, is robust and the eventual outcome does not significantly depend on the environmental changes that brought it about in the first place. Geography plays an important role. Somewhat paradoxically, a certain extent of geographical cohesion is needed for the build-up of a global social network. But if such cohesion is too strong, it is detrimental to the whole process and can block it altogether. Hence some intermediate level of cohesion is optimal. In essence, globalization is both the outcome and the consequence of a dense and diverse pattern of distant connections. This opens up the possibility of multiple equilibrium configurations. That is, due to history, chance, or policy some economies may have proven successful while others facing the same environmental circumstances have not. In fact, we shall see that suitable policy measures, impinging temporarily on the economy, may explain persistently different outcomes.

The rest of the paper is organized as follows. Section 2 reviews some related literature and briefly outlines a companion paper where we test empirically some basic implications of our theory. Section 3 presents the model: in Subsection 3.1 we describe the interaction framework, while in Subsection 3.2 the dynamics. Section 4 carries out the analysis, decomposing it in three parts. Firstly, Subsection 4.1 succinctly discusses some numerical simulations that illustrate important features of the model. Secondly, Subsection 4.2 undertakes the theoretical analysis in a simplified benchmark setup applying to a limit infinite-population context. Thirdly, Subsection 4.3 extends the theory to a general context with finite populations and any parameter configuration. Section 5 concludes the main body of the paper with a summary of its content and an outline of future research. For the sake of smooth exposition, formal proofs and other details of our analysis are relegated to the Appendix.

2 Related literature

As advanced, the bulk of economic research on the phenomenon of globalization has been of an empirical nature, with only a few papers addressing the issue from a theoretical perspective. Two interesting theoretical papers that display a certain parallelism with our approach are Dixit (2003) and Tabellini (2008). In both of them, agents are distributed over some underlying space, with a tension arising between the advantages of interacting with far-away agents and the limits to this imposed by geographical distance. Next, we outline these papers and contrast them with our approach.

The model proposed by Dixit (2003) can be summarized as follows: (i) agents are arranged uniformly on a ring and are matched independently on each of two periods; (ii) the probability that two agents are matched decreases with their ring distance; (iii) gains from matching (say trade) grow with ring distance; (iv) agents' interaction is modelled as a Prisoner's Dilemma; (v) information on how any agent has behaved in the first period arrives at any other point in the ring with a probability that decays with distance.

In the context outlined, one obtains the intuitive conclusion that trade materializes only between agents that do not lie too far apart. Trade, in other words, is limited by distance. To overcome this limitation, Dixit contemplates the operation of some "external enforcement." The role of it is to convey information on the misbehavior of any agent to *every* potential future trader, irrespective of distance. Then, under the assumption that such external enforcement is quite costly, it follows that its implementation is justified only if the economy is large. For, in this case, the available gains from trade are also large and thus offset the implementation cost.

The second paper, Tabellini (2008), relies on a spatial framework analogous that of Dixit (2003). In it, however, distance bears *solely* on agents' preferences: each matched pair again plays a modified Prisoner's Dilemma, but with a warm-glow altruistic component in payoffs whose size falls with the distance to the partner. Every individual plays the game only *once*. This allows the analysis to dispense with the information-spreading assumption of Dixit's model, which presumes that agents are involved in repeated interaction. Instead, the distinguishing characteristic of Tabellini's model is that agents' preferences (specifically, the rate at which the warm-glow component decreases with distance) are shaped by a process of intergenerational socialization à la Bisin and Verdier (2001).

In a certain sense, altruist preferences and cooperative behavior act as strategic complements in Tabellini's model. This, in turn, leads to interesting coevolving dynamics of preferences and behavior. For example, even if both altruism and cooperation start at low levels, they can reinforce each other and eventually lead the economy to a state with a large fraction of cooperating altruists – i.e. agents who care for, and cooperate with, even relatively far-away partners. Under reasonable assumptions, such steady state happens to be unique. There are, however, interesting variants of the setup where the enforcement of cooperation (amounting to the detection of cheating and the offsetting of its consequences) is the endogenous outcome of a political equilibrium, and this allows for multiple steady states that depend on initial conditions.

In resemblance with the two papers just summarized, our approach also attributes to some suitable notion of "geographical" distance a key role in shaping the social dynamics. In Dixit (2003) and Tabellini (2008), however, the impact of such *exogenous* distance is direct: the ability to sustain cooperation (on the basis of either observability or altruism) is taken to decrease in it. In our case, instead, the relevant distance in the establishment of new partnerships is *social* and *endogenous*, for it is determined by the evolving social network. It is precisely through the evolution of the social network that geographic distance plays a *indirect* role in the model. Geographically closer agents are assumed to enjoy a higher arrival rate of collaboration opportunities, although these opportunities in effect materialize only if their *current* social distance is short enough.[1]

Next, let me turn to the empirical literature concerned with the phenomenon of globalization. Typically, it has focused on a single dimension of the problem, such as trade (Dollar and Kraay (2001)), direct investment (Borensztein *et al.* (1998)) or portfolio holdings (Lane and Milesi-Ferretti (2001)). A good discussion of the conceptual and methodological issues to be faced in developing coherent measures along different such dimensions are systematically summarized in a handbook prepared by the OECD (2005 *a*,*b*). But, given the manifold richness of the phenomenon, substantial effort has also been devoted to developing composite indices that reflect not only economic considerations, but also social, cultural, or political. Good examples of this endeavour are illustrated by the interesting work of Dreher (2006) –see also Dreher *et al.* (2008) – or the elaborate globalization indices periodically constructed by A.T. Kearney/Foreign Policy (2006) and the Centre for the Study of Globalization and Regionalization (2004) at Warwick.

These empirical pursuits, however, stand in contrast with our approach in that they are not conceived as truly systemic. That is, the postulated measures of globalization are based on a description of the individual characteristics of the different "agents" rather than on how they interplay within the overall structure of interaction. Our model, instead, calls for systemic, network-like, measures of globalization. A few papers in the recent empirical literature that move in this direction are Kali and Reyes (2007), Arribas *et al.* (2009), and Fagiolo *et al.* (2010). They all focus on international trade flows and report some of the features of the induced network that, heuristically, would seem appealing, e.g. clustering, node centrality, multi-step indirect flows, or internode correlations. Their objective is mostly descriptive, although Kali and Reyes show that some of those network measures have a significant positive effect on growth rates when added to the customary growth regressions. These papers represent an interesting first attempt to bring genuinely global (network) considerations into the discussion of globalization. To make the exercise truly fruitful, however, we need some explicitly formulated theory that guides both the questions to be asked as well as the measures to be judged relevant

In a companion empirical paper, Duernecker, Meyer, and Vega-Redondo (2011) have built on the theory presented here to undertake a step in this direction. First, it introduces an operational counterpart of the measure of integration that is based on our present theoretical model. Then, it checks whether that measure is a significant variable in explaining intercountry differences in growth performance. In this exercise we start by identifying systematically the control variables that are to be considered and then, most crucially, address the key endogeneity problem that lies at the core of the empirical issue.[2] We find that our measure of integration is a robust and very significant explanatory variable that supersedes traditional measures of openness (e.g. the ratio of exports and imports to GDP), rendering them statistically insignificant. This suggests that the network-based approach to integration that is proposed here adds a systemic perspective that is rich and novel. We refer the interested reader to the aforementioned companion paper for details.

[1]As we shall explain, the only exception is when the two agents are immediate geographic neighbors

[2]The selection of control variables is conducted through a large-scale use of the so-called Bayesian averaging methods used in the literature (see e.g. Sala-i-Martin *et al* (2004)). The issues raised by endogeneity, on the other hand, are tackled by relying on the use of instrumental-variable methods suitable panel data.

3 The model

The formal presentation of the model is divided in two parts. First, we describe the underlying (fixed) spatial setup and the (changing) social network that is superimposed on it. Second, we specify the dynamics through which the social network evolves over time from the interplay of link creation (innovation) and link destruction (volatility). The formulation of the model is kept at an abstract level in order to stress its versatility. In Appendix A, however, we spell out in detail a concrete setup that provides a game-theoretic foundation for it. In this setup, agents are assumed to be involved in a collection of repeated games of uncertain duration, with cooperation being supported (at an equilibrium of the underlying population game) by both bilateral and third-party threats of punishment.

3.1 The basic setup

Let N be a fixed (large) population of n agents, evenly spread along a one-dimensional ring of fixed length. To fix ideas, we shall speak of this ring as reflecting physical space but, as is standard, it could also embody any other relevant characteristic (say, ethnic background or professional training). The location of each individual in the ring is assumed fixed throughout. For any two agents i and j, the "geographical" distance between them is denoted by $d(i, j)$. By normalizing the distance between two adjacent agents to one, we may simply identify $d(i, j)$ with the minimum number of agents that lie between i and j along the ring, including one of the endpoints.

Time is modelled continuously. At each point in time $t \geq 0$, there is social network in place, $g(t) \subset \{ij \equiv ji : i, j \in N\}$, each of its links interpreted as an ongoing project run in collaboration by the two agents involved. This introduces an additional notion of distance between agents – their social (or network) distance – that is given by the length of the shortest network path connecting any two nodes. (If no such path exists, their social distance is taken to be infinite.) In general, of course, the prevailing social distance $\delta_{g(t)}(i, j)$ between any two nodes i and j can be higher or shorter than their geographical distance $d(i, j)$; see Figure 1 for an illustration.

3.2 Dynamics

The law of motion of the system identifies the current state at any given t with the prevailing network $g(t)$. The state changes due to two forces alone: *innovation* and *volatility.* The first one leads to the creation of new links, while the second one involves the destruction of existing ones. We describe each of them in turn.

3.2.1 Innovation

At every t, each agent i gets an idea for a new project at a fixed rate $\eta > 0$. This project, however, requires the collaboration of one other particular agent j. From an *ex ante* viewpoint, the probability $p_i(j)$ that the agent required by i's idea is any given j is taken to satisfy:

$$p_i(j) \propto 1/\left[d(i, j)\right]^{\alpha}, \tag{1}$$

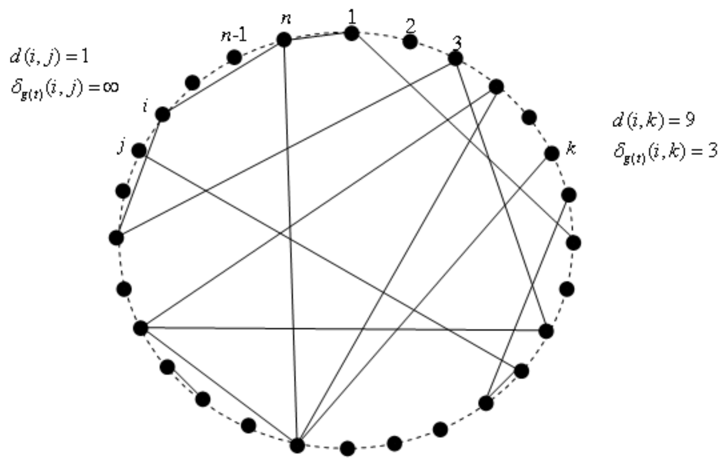


Figure 1: Snapshot of a situation at some t. By way of illustration, note that whereas the social distance is maximal (infinity) between agents i and j who are neighbors on the ring (i.e. their geodistance is the minimal of 1), the same comparison for i and k yields the opposite conclusion, namely, their geodistance is higher than their social distance.

for some $\alpha > 0$, i.e. it decays with the geographical distance (geodistance, for short) between i and j at the rate α. Since this abstract formulation admits a wide variety of concrete interpretations, in order to fix ideas we shall discuss just two of them. The first is the most simplistic one. It conceives i and j as actually meeting physically at the time i has the idea, thus making j the feasible partner to carry it out. In this interpretation, decay just reflects the fact that closer agents meet more often than distant ones. A second alternative interpretation is based on the notion that fruitful collaboration not only requires not only complementarity of skills between the agents involved but also compatibility in a number of dimensions (e.g. language, norms, expectations, etc.). In this case, geographical decay can be thought as embodying the idea that distant agents share less of the common ground on which such a compatibility is based upon. Hence, *a priori*, it is more likely that the partner required lies geographically close rather than far away. In any case, whatever the interpretation, it is the exponent $\alpha > 0$ that modulates the strength of these considerations. For the sake of conciseness we shall refer to this parameter as (geographical) *cohesion*.

Consider then an agent i who has an idea for a project that needs the concourse of agent j. When will this idea indeed materialize into an actually running project? Our model prescribes that this will happen if, and only if, the following two conditions are *jointly* met:

L1 Agents i and j are not already involved in an ongoing project together.

L2 These agents are close, *either* geographically or socially, as captured by two *positive* parameters, ν and μ respectively. More precisely, we require that at least one of the following conditions holds:

(a) they are at a geodistance no higher than ν;

(b) they are at a social distance no higher than μ.

Condition L1 captures in a very simple and stark way a key idea of our setup – namely, that there is a limit to (or saturation of) the fruitful opportunities that can be carried out by repeatedly relying on the same partners. This can be motivated by the assumption that the economic potential of any given combination of skills is restricted to a "niche" of limited size. Thus, in order for an agent to expand her set of economic opportunities, she cannot just rely on the intensive margin but needs to exploit the "extensive margin."[3]

On the other hand, Condition L2 specifies the restrictions that must be satisfied if a (non-redundant) opportunity to form a link is to materialize. It requires that the two agents involved not be too far away (as given by ν and μ) in *at least one* of the two notions of distance considered. The specific value of ν has just minor implications on the analysis since it concerns the exogenously fixed notion of distance associated to "geography." (To be sure, the fact that ν is positive plays a crucial role, as we explain below.) We simplify matters, therefore, and choose $\nu = 1$ throughout, thus particularizing L2(a) to the mere condition that agents i and j be geographic neighbors.

Instead, concerning social distance, the precise value of $\mu \in \mathbb{N}$ considered in L2(b) does have interesting implications on the analysis, so we keep it as an explicit parameter of the model. Conceptually, it reflects the range at which the social network can operate as an *effective channel* for information, access, or trust. In principle, there could be many alternative mechanisms at work in this respect, so we choose to refer to μ very generally as the (quality of the) *institutions* prevailing in the economy. Within the microfounded scenario described in Appendix A, this parameter will be identified with the so-called "circle of trust" – i.e., roughly, the maximum social distance that may separate two individuals such that either of them is ready to punish any third party who has behaved uncooperatively with the other one. [4]

3.2.2 Volatility

As advanced, project volatility – i.e. link decay – is the second force governing the dynamic process. We choose to model it in the simplest possible manner. Specifically, we posit that every existing link/project is discontinued (say, because it becomes "obsolete") at a constant rate λ, which is normalized to unity ($\lambda = 1$) without loss of generality. Link destruction is therefore modelled as an exogenous process, letting the interplay between the social network and the overall dynamics be fully channelled through the mechanism of link formation alone.[5]

---

[3]Given the assumed ex-ante homogeneity across agents, we posit that the same strict limit of at most one standing project applies to every pair of agents. Allowing for some inter-agent heterogeneity in this respect could have interesting implications on the process of network formation, and would be in line with the substantial skewedness in the distribution of innovation that is observed in the real world – see, for example, Jaffe and Trajtenberg (2002).

[4]At a heuristic level, μ can be seen as modulating the extent to which the population abides by the cooperative norm embodied by the underlying equilibrium. Such a game-theoretic interpretation is in line with the influential work of Coleman (1988, 1990), who stressed that the cohesiveness (or, as he called it, "closure") of the social network is often key in deterring opportunistic behavior. This is also a theme that has been revisited extensively by recent literature, also casting the problem in game-theoretic terms – see e.g. Greif (1993), Haag and Lagunoff (2006), Lippert and Spagnolo (2006), Vega-Redondo (2006), Karlan *et al.* (2009) and Jackson *et al.* (2010).

[5]Invoking considerations of trust or monitoring analogous to those that have been suggested for link formation, it could be postulated, for example, that the rate at which every link disappears grows with the network distance of the agents involved (excluding their own link). We conjecture that nothing essential in the analysis would be affected by this modification.

4 Analysis

In a nutshell, the network formation process modelled in the preceding section can be succinctly described as the *struggle of innovation against volatility.* The primary objective of our analysis is to understand the conditions under which such a struggle allows for the maintenance, in the long run, of a high level of economic interaction (i.e. connectivity). More specifically, our focus will be on how the long-run performance of the system is affected by the (sole) three parameters of the model: η (the rate of innovation), α (geographical cohesion) and μ (institutions).

The discussion in this section is organized in three parts. First, we start with some numerical simulations that are useful to highlight issues and insights that will arise in our general analysis. Second, we develop a formal theory that is directly applicable only to the limit case of an infinite population, but nevertheless sheds light on the key mechanisms at work in the general setup. Third, we devise a numerical approach to solving the model that allows us to generate a full array of comparative-statics results for all parameter configurations and thus provides a detailed understanding of the model.

4.1 Some preliminary numerical simulations

By discretizing in a natural way the continuous-time theoretical model, we have conducted numerical simulations that provide us with some informal but useful illustration of the range of behavior displayed by the system in the long run (see Appendix C for the algorithm and other details of the numerical procedure). For the sake of focus, our main objective at this point will center on the following two issues. First, we want to obtain an intuitive grasp of what changes might be expected in the economy when it turns global and what are the main mechanisms underlying this transformation. Second, we would like to obtain some preliminary insights on the role played by geography as well as the positive and negative effects of geographical cohesion.

Consider first the simulation results described in Figure 2. Each of its four diagrams traces different loci of steady-state values for four different variables (measured on the vertical axes) associated to different values of the innovation rate η (measured on the horizontal axes) and fixed values of the remaining parameters. The main point to make on the parameters considered here is that the level of geographical cohesion is relatively low, $\alpha = 0.5$. This value, in particular, is well below the threshold of unity that, as we shall see (cf. Subsection 4.2), marks a qualitative change in the environment.

The steady-state variables recorded in each of the diagrams of Figure 2 are as follows:

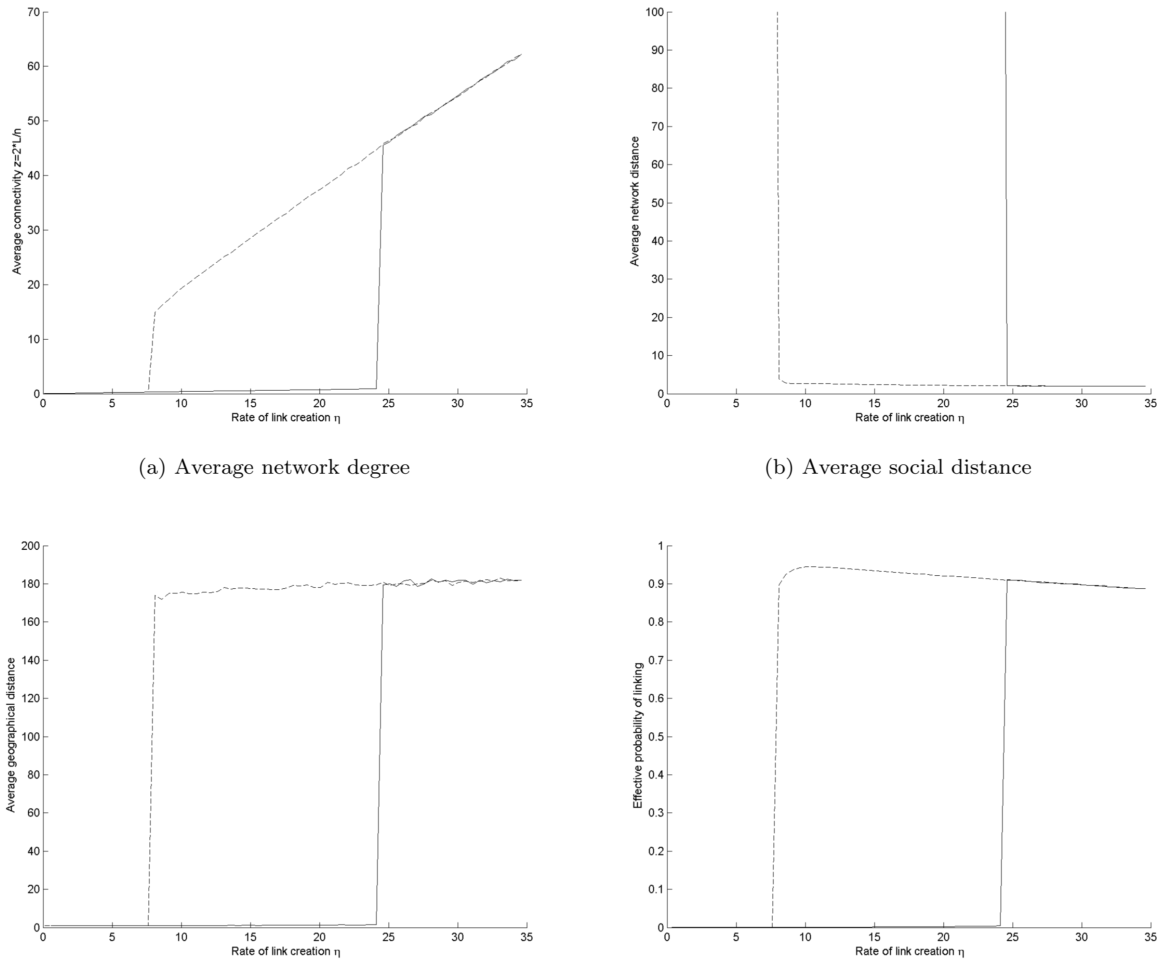
(a) average degree, as given by $\frac{2L}{n}$, where L is the total number of links and n is the population size;

(b) average social distance among agents, computed as $\frac{1}{n(n-1)} \sum_{i,j \in N} \delta(i,j)$;

(c) the average geographical distance across links, obtained as $\frac{1}{L} \sum_{ij \in g} d(i,j)$;

(d) the fraction of agent pairs who, upon receiving a link-creation opportunity, actually form a new link.

In each of the four diagrams there are two loci of steady-state values, corresponding to the only two (stable) steady-state aggregate configurations that arise: one corresponding to initial conditions where the social



(a) Average network degree

(b) Average social distance

(c) Average geodistance

(d) Effective linking probability

Figure 2: Discontinuous transition in network connectivity, and hysteresis, as the innovation rate η changes for given (low) $\alpha = 0.5$ and $\mu = 3$ (with $\lambda = 1$, $n = 1000$).

network displays a high connectivity (dashed line) and another where the initial social network is sparse (solid line).

Let us first focus on Diagram (a) in Figure 2. The two steady-state loci displayed in it (solid and dashed lines) display a similar pattern. Initially, there is a low-connectivity regime where not only the average degree is low but the responsiveness to increases in η is very low as well. Then, there is an abrupt transition to a high connectivity regime where both the average degree and the responsiveness to increases in η becomes high. Overall, we find that the parameter space for η can be partitioned into three different regions: a low region where every η induces a unique steady-state outcome with low connectivity; a middle region where steady states with either high and low connectivity can be reached, depending on initial conditions; and a high region where again, independently of initial conditions, the system converges to a single steady-state outcome with high connectivity. An interesting consequence of this behavior is that, if one traces the effect of gradual changes in η on the steady state the system displays strong *hysteresis*. That is, around the points where regime transitions occur, the responses to increases or decreases in η are *not* symmetric. [6]

To understand the mechanism underlying the pattern just described, it is useful to turn to Diagrams (b) and (c) in Figure 2, which are the exact counterparts of the previous Diagram (a) but now pertaining to average social distance (per agent pair) and average geodistance (per existing link). We find that, corresponding to each transition to (or from) a high-connectivity regime there is a corresponding transformation of the social network that, in our terminology (recall the Introduction), has to be understood as a transition to (or from) globalization. In other words, such network transitions may be seen as the other side of the coin of those in terms of connectivity, i.e. they must go hand in hand. Why is this the case?

To answer that question, we focus on Diagram (d), which shows what is the main implication of globalization: it induces a sharp increase in the effective linking probability. Only because there is an abrupt fall in social distances (the world becomes "small") link creation becomes so much more effective. And this increased effectiveness is crucially needed to support a high-connectivity configuration as a steady state. For, in this case, many links are continuously being destroyed and hence many need to be continuously created as well to arrive at a stationary situation.

As indicated, the behavior displayed in Figure 2 is obtained in a scenario where the level of geographical cohesion is relatively low ($\alpha = 0.5$) and, in particular, below one. What happens if it is significantly higher? This is explored in the simulations described in Figure 3, where the level of cohesion is raised to a value of $\alpha = 2$ (i.e. above the threshold of one). For the sake of a clear comparison, the remaining parameters are exactly the same as before, and so is the interpretation of the curves displayed in the different diagrams.

Figure 3 reveals a striking contrast with the low-cohesion scenarios. Specifically, the following key differences are observed in the present high-cohesion context:

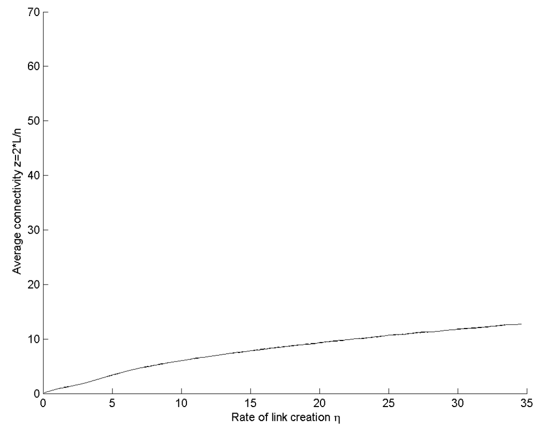
(i) For any given η, the system converges to a unique steady-state outcome, independently of initial conditions.

(ii) The responsiveness to η is always significant, even at the lowest levels when the underlying network is very sparse.

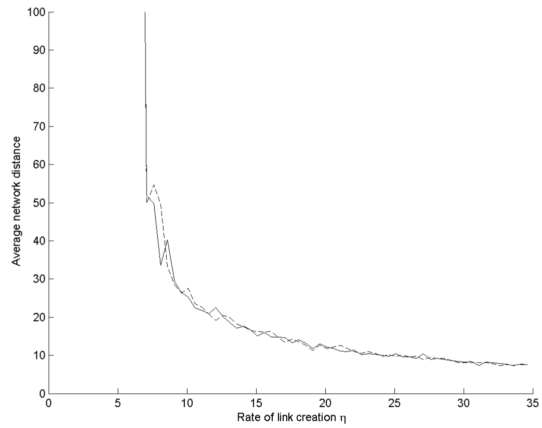
(iii) The effective linking probability remains well below unity, decreasing in η for most of its range.

Thus the first interesting observation, stated in (i) above, is that initial conditions do not matter in the long run if cohesion is relatively high. Thus, even if the process starts from a very sparse network, the steady-state configuration to which the process converges is not affected by such a "difficult start," even if η is relatively low. The key reason for this has to do with the significant structure on the matching process induced by high geographical cohesion. Under high cohesion, potential partners tend to be geographically close, which in turn implies that it is likely that both have been previously matched (and formed a link) with some common third agents who are geographically close as well. In the end, this implies that a short network path is likely to exist between potential partners (i.e. meeting agents), who can then form a link if they do not yet have

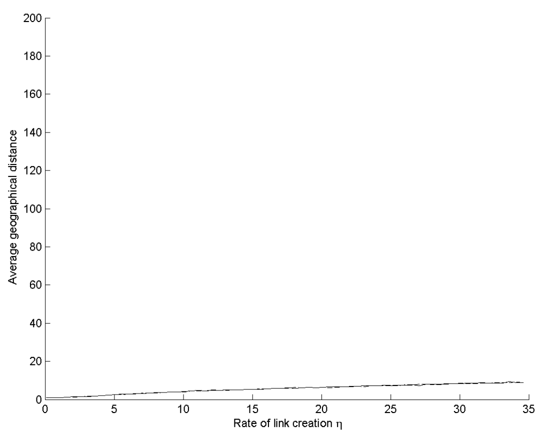
[6] This, of course, is simply a reflection of the dependence of initial conditions that occurs in the middle η region. The implications of it for policy are potentially important, as briefly outlined in Section 5.



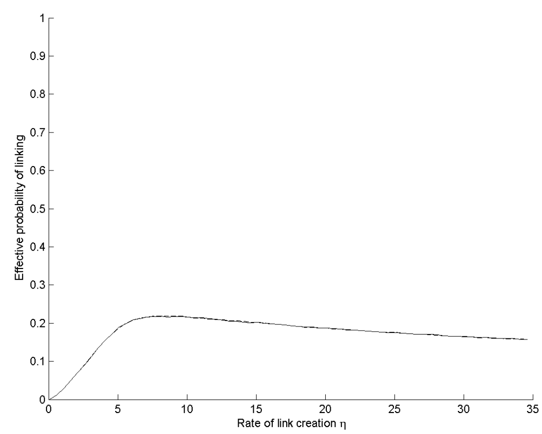
(a) Average network degree



(b) Average social distance



(c) Average geodistance



(d) Effective linking probability

Figure 3: Continuous change in network connectivity, without hysteresis, as the innovation η rate changes when $\alpha = 2$ and $\mu = 3$ ($\lambda = 1$, $n = 1000$).

one (recall L1-L2 in Subsection 3.2.1). Initial conditions, therefore, are no longer relevant to predict where the process will end up in the long run, and only the value of the underlying parameters matters.

Essentially the same idea underlies point (ii) above. That is, even if η is low, high geographic cohesion allows increases in the innovation rate η to translate into increases in connectivity of a comparable magnitude. But then we are drawn upon an important trade-off, which is highlighted by item (iii) above: higher cohesion not only yields linking benefits, but also comes at a cost. Congestion, specifically, turns into an important concern. It reflects the saturation of the intensive margin in link creation that results from the impossibility of having more than one link per agent pair. This problem, in fact, is not only manifested in the lower effectiveness of link creation stated in (iii). It is also evident from a comparison of Diagrams (a) in both Figures 2 and 3. Such a comparison shows that, for high values of the innovation rate, the level of connectivity achieved under $\alpha = 0.5$ is substantially higher than that obtained for $\alpha = 2$ – the only proviso is that η be high enough that globalization has materialized in the first case.

Overall, the former discussion suggests that there should be an "optimal level of cohesion" which strikes the best balance between the need to have some geographic structure to build the social network from scratch and the detrimental congestion effects caused by too much cohesion. Indeed, our formal analysis will show that such a trade-off is a general feature of the model, and that the induced tension can be resolved in a natural way, as a function of the remaining parameters. For example, we shall find that the higher is the quality of institutions or the faster innovation opportunities arrive, the lower is the (optimal) level of cohesion that the economy needs to attain its maximum long-run potential.

4.2 Benchmark theory

In this section, we study a benchmark limit context where the population is very large ($n \to \infty$) and the underlying network is parametrized by its average degree. The insights gained from this scenario will help us understand some of the key forces at work in the general model – in particular, those that were explained in a merely intuitive manner in the preceding subsection. The theory will provide, for example, the prediction that there is a threshold on geographical cohesion ($\alpha = 1$ in the limit scenario) that marks the point above which the economy is able to build up a dense social network. Another theoretical prediction will be that, in those cases where cohesion is not too strong, the transition towards globalization must be sharp (i.e. discontinuous) and locally irreversible (display hysteresis) as η changes. These predictions are of course very much in line with the simulation results discussed in Subsection 4.1 for the lower value of α.

Our analysis will revolve around the following simple characterization of the stable steady states of the process. Let ϕ denote the conditional linking probability prevailing at some steady state – i.e. the probability that a randomly selected agent who receives an innovation draw effectively succeeds in forming a new link. Then, the induced expected rate of project/link creation is simply the product $\phi\eta n$ where η is the innovation rate and n the population size (for the moment, taken to be finite). On the other hand, if we continue to denote the average degree (average number of links per node) by z, then the expected rate of project/link destruction is $\lambda(z/2)n = \frac{1}{2}zn$, where recall that we are making $\lambda = 1$ by normalization.[7] Thus, finally, bringing the former two expressions together and cancelling n on both sides, the condition "rate of link creation = rate of link destruction" that characterizes a steady state can be written as follows:

$$\phi\,\eta = \frac{1}{2}z. \tag{2}$$

Naturally, the difficulty here lies in a proper determination of ϕ, which is an *endogenous* variable that depends on the state of the system, i.e. on the prevailing network $g(t)$. To address the problem, we follow an approach that is common in the modern theory of complex networks. It rests on the assumption that $g(t)$ is a randomly generated network, and the induced distribution over possible networks is suitably parametrized by the average network degree $z(t)$.[8] This allows us to postulate the existence of a function $\Phi : \mathbb{R}_+ \to [0, 1]$

[7] Also bear in mind that the number of links is *half* the total degree because every link contributes to the degree of its two nodes.

[8] This is, for example, the approach undertaken by two of the canonical models in the random network literature: the Erdös-Rényi networks (Erdös and Rényi (1959)) and the Scale-Free networks (Barabási and R. Albert (1999)). Even though

specifying the conditional linking probability $\Phi(z)$ associated to each average degree z, and hence write the steady-state condition (2) as follows:

$$\Phi(z) = \frac{1}{2\eta} z. \tag{3}$$

The above condition may be conceived as an equation in z, from which the equilibrium network connectivity prevailing in the steady states of the process can, in principle, be suitably determined.

As advanced, the properties of the resulting steady states crucially depend on the value of α. We shall need to consider, specifically, two qualitatively different regions:

- **Low geographical cohesion** (LGC): $\alpha \leq 1$
- **High geographical cohesion** (HGC): $\alpha > 1$

In the first LGC region, "geography" (in our general, and possibly metaphorical, sense of the word) has little bearing on how new ideas arise; in the second one, HGC, the opposite applies. We shall show that each of these two regions in the parameter space display qualitatively different implications – not only for the set of configurations that qualify as possible equilibria but, most crucially, on their respective stability.

A central role in the analysis is played by the probability that the collaborator needed by any particular agent in order to undertake a new project is precisely one of his two neighbors on the ring. From (1), this probability is simply given by the inverse of the following expression:

$$\zeta(\alpha, n) \equiv \sum_{d=1}^{(n-1)/2} \frac{1}{d^\alpha}, \tag{4}$$

where, for simplicity, we assume that n is odd. In our study of the limit case $n \to \infty$, we are led to the following magnitude:

$$\zeta(\alpha) \equiv \lim_{n\to\infty} \zeta(\alpha, n) = \sum_{d=1}^{\infty} \frac{1}{d^\alpha}, \tag{5}$$

which is known as the real-valued *Riemann Zeta Function*. It is is a standard result[9] that

$$\zeta(\alpha) < \infty \quad \Longleftrightarrow \quad \alpha > 1. \tag{6}$$

Thus, if $\alpha > 1$, the aforementioned probability that an agent meets one of his geographic neighbors is positive, given by: $[\zeta(\alpha, n)]^{-1}$. Instead, if $\alpha \leq 1$, the series in (5) diverges and we shall interpret $[\zeta(\alpha)]^{-1}$ to be equal to zero.

Our analysis starts with the following simple result that establishes useful properties on the linking probability ϕ prevailing in a finite network where the population size is n. (The proof can be found in Appendix B.)

---

each of them displays polar features in a number of key respects (e.g. in terms of the variance displayed by the induced degree distributions), in both cases the family of networks under consideration can be parametrized by the average degree z (cf. Vega-Redondo (2007)).

[9] See e.g. Apostol (1974, p. 192).

Proposition 1 *Let $g(t)$ be the network prevailing at some time t in a population of size n, and consider any given agent i who is enjoying an innovation (link-creation) opportunity. Denote by $\phi_i(t)$ the conditional probability that it actually establishes a new link (under the link-formation rules L1-L2 in Subsection 3.2.1) and let $M_i(g(t))$ stand for the size of the component to which it belongs.*[10] *Then, for* any *randomly selected agent i, we have:*

$$\phi_i(t) \leq \frac{M_i(g(t)) + 1}{2} \left[\zeta(\alpha, n)\right]^{-1}. \tag{7}$$

In addition, if agent i happens to be isolated *(i.e. has no partners), then*

$$\phi_i(t) \geq \left[\zeta(\alpha, n)\right]^{-1}. \tag{8}$$

Let us now return to the assumption that the network prevailing at any t can be conceived as a large random network suitably parametrized by its average degree z. Then, it is well known that many interesting (statistical) properties of the network depend on z in a "threshold" manner.[11] For our purposes, an important such property concerns the expected component size associated to a randomly selected node, a magnitude which will be denoted by $\bar{M}(z;n)$ in order to reflect its dependence on z and the (finite) population size n. The aforementioned threshold property can be formulated as follows (cf. Durrett (2007, Ch. 3)):

$$\exists \hat{z} > 0 \quad \text{s. t.} \quad \lim_{n\to\infty} \bar{M}(z;n) < \infty \quad \Longleftrightarrow \quad z < \hat{z}. \tag{9}$$

That is, the average component size turns from finite to unbounded as the average degree exceeds a particular threshold.

Extending previous notation, let us write the function in (3) that specifies conditional linking probabilities as $\Phi(z;\, \alpha)$, in order to make explicit its dependence on α. Based on (9), we now formulate some derived assumptions on this function that capture what must be its asymptotic behavior for large populations. Specifically, we posit that its local behavior around $z = 0$ under the two polar scenarios considered before – i.e. LGC ($\alpha \leq 1$) and HGC ($\alpha > 1$) – is respectively given by the following two contrasting implications.

A1. Let $\alpha \leq 1$. Then, there exists some $\hat{z} > 0$ such that for all $z \leq \hat{z}$, $\Phi(z;\, \alpha) = 0$.

A2. Let $\alpha > 1$. Then, $\Phi(0;\, \alpha) = [\zeta(\alpha)]^{-1} > 0$

Assumption A1 applies to the LGC scenario and simply follows from (6), (7), and (9) if one interprets $\Phi(z,\, \alpha)$ as the limit of the linking probability of a randomly selected node i when $n \to \infty$. In contrast, A2 applies to the HGC scenario and follows from (6), (7), and (8) as $n \to \infty$, since when $z = 0$ almost all nodes must then be isolated.

Finally, we also postulate two further assumptions that are largely of a technical nature:

A3. The function $\Phi(z;\, \alpha)$ is jointly continuous in z and α.

[10] As usual, a component is defined to be a *maximal* set of nodes such that every pair of them are connected, either directly by a link or indirectly through a longer path. The size of a component is simply the number of nodes included in it

[11] See, for example, the extense survey of Newman (2003) or the very accessible monograph by Durrett (2007) for a good presentation of the different results from the theory of random networks that will be used here. For an exhaustive and rigorous account of this field of research, see the classic book by Bollobás (1985) in its latest edition (2001).

A4. Given any α, there exists some $\tilde{z} \geq 0$ such that $\Phi(\tilde{z};\, \alpha) > 0$. Moreover, for any z', z'' such that $z'' > z'$, $\Phi(z';\, \alpha) > 0 \Rightarrow \Phi(z'';\, \alpha) > 0$.

Assumption A3 is a standard regularity condition that facilitates the analysis. On the other hand, A4 comes in two parts. The first part states that there is a sufficiently high level of network connectivity for which the linking probability of agents is positive. This implicitly presumes that the institutional parameter μ – which has been kept in the background in the present analysis – is high enough. An elaboration on this point is carried out in Remark 1 below, while a full-fledged analysis of the role of this parameter is postponed to the study of the finite-population setup that is undertaken in Section 4.3. The second part of A4 states that, once a sufficient level of connectivity is attained such that the linking probability is positive, further increases in it still preserve a positive (but possibly decreasing) probability. Intuitively, this captures the idea that, in a context where the population is arbitrarily large and the average node degree is finite, linking probabilities cannot fall down to zero just because of saturation.

Conceiving assumptions A1-A4 as an appropriate description of how the system behaves for very large populations, we now contemplate the following dynamics for its aggregate state (which, as explained, is embodied by the average degree z).

$$\dot{z} = \eta\, \Phi(z;\, \alpha) - \frac{z}{2} \tag{10}$$

The above differential equation simply reflects the dynamics through which the net effect of link creation and link destruction changes the average degree of the network. Our ensuing results highlight some of the main conclusions that follow from the analysis of (10). As before, all proofs can be found in Appendix B.

We start with two results that establish the key role played by geographical cohesion in determining whether the economy is capable, or not, of building up a significant social network "from scratch."

Proposition **2** *Let $\alpha \leq 1$. Then, the state $z = 0$ is asymptotically stable.*[12]

Given any initial condition $z_0 \geq 0$, denote by $[\varpi(z_0, t)]_{t\geq 0}$ the trajectory induced by (10) that starts at z_0.

Proposition **3** *Let α >1. Then, if $\eta > 0$, the state $z = 0$ is* not *asymptotically stable. Furthermore, there exists a* unique *steady state $z^* > 0$ that satisfies the following condition:*

$$\exists \epsilon > 0 \quad s.t. \quad \forall z(0) \leq \epsilon, \quad \varpi\,(z(0), t) \to z^*. \tag{11}$$

The former two results stress the importance of cohesion in the build-up of the social network. When cohesion is low, a virtually empty network with $z = 0$ is robust (i.e. locally stable). Instead, if the cohesion parameter α exceeds the critical threshold of one, such a configuration becomes unstable and, in fact, there is a unique positive average degree to which the economy converges from any sufficiently small extent of overall connectivity.

The key ideas underlying the previous two propositions are illustrated in Figures 4 and 5, where steady states are represented by intersections of the function $\Phi(z)$ and the ray of slope $1/(2\eta)$. First, Figure 4

[12] As standard, a state is said to be asymptotically stable if any trajectory of the system hat starts sufficiently close to it converges to that state. This notion, therefore, is one of *local* stability.

depicts a situation with low cohesion (i.e. $\alpha < 1$), which implies (from A1) that $\Phi(z)$ displays a uniform value of zero up to some average degree $\hat{z} > 0$. Then we have that, if η is small enough (e.g. $\eta = \eta'$ in this figure), the induced ray is so steep that the only steady state involves a network displaying a zero average degree. Instead, for larger values of η such as η'', there are two *additional* steady states with a positive average degree (z_2 and z_3), the first being unstable and the second stable. The key point to stress here, however, is that, for *any* value of η, the zero-degree equilibrium is *locally stable* – so, for example, in the situation illustrated in Figure 4, this happens both for the low η' as well as for the high η''.

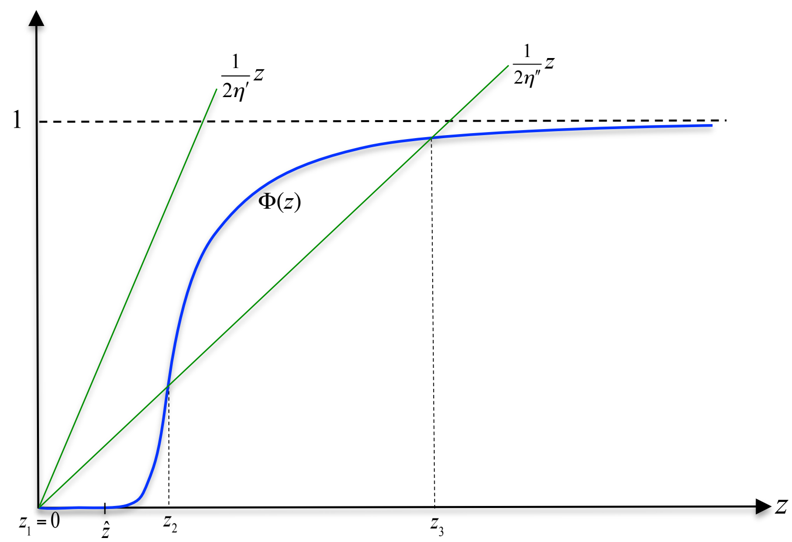


Figure 4: The effect of changes in η on the equilibrium degree under low cohesion ($\alpha < 1$)). See the text for an explanation.

In contrast, Figure 5 depicts a situation corresponding to a cohesion parameter $\alpha > 1$. In this case, $\Phi(0) = [\zeta(\alpha, n)]^{-1} > 0$, which implies that every value of η induces a positive equilibrium degree, to which the system converges from an initially empty network (i.e. one with $z = 0$). Provided η is high enough, such an equilibrium degree is low. If the innovation rate increases, the average degree continues being low as long as η remains under a certain threshold ($\hat{\eta}$ in the figure). But as η just grows beyond this threshold, a discontinuous (possibly very large) increase in the equilibrium degree takes place (from z_4 to z_5 in the example considered). Thereafter, if η increases even further, the degree displayed by the corresponding equilibrium grows unboundedly (e.g. up to z_5 and beyond). As we show below, the pattern described above always occurs when the economy displays a level of cohesion that is slightly above the amount required to escape the empty network – i.e. the value of α is above one but not too large. Instead, the unboundedness of the degree as a function of η applies generally for all values of $\alpha > 1$.

PROPOSITION **4** *Given any $\alpha > 1$, denote by $z^*(\eta)$ the steady state that is established in Proposition 3. The function $z^*(\cdot)$ is strictly increasing and unbounded (i.e. for any $\bar{z}$, there exists some $\bar{\eta}$ such that if $\eta \geq \bar{\eta}$, then $z^*(\eta) \geq \bar{z}$.*

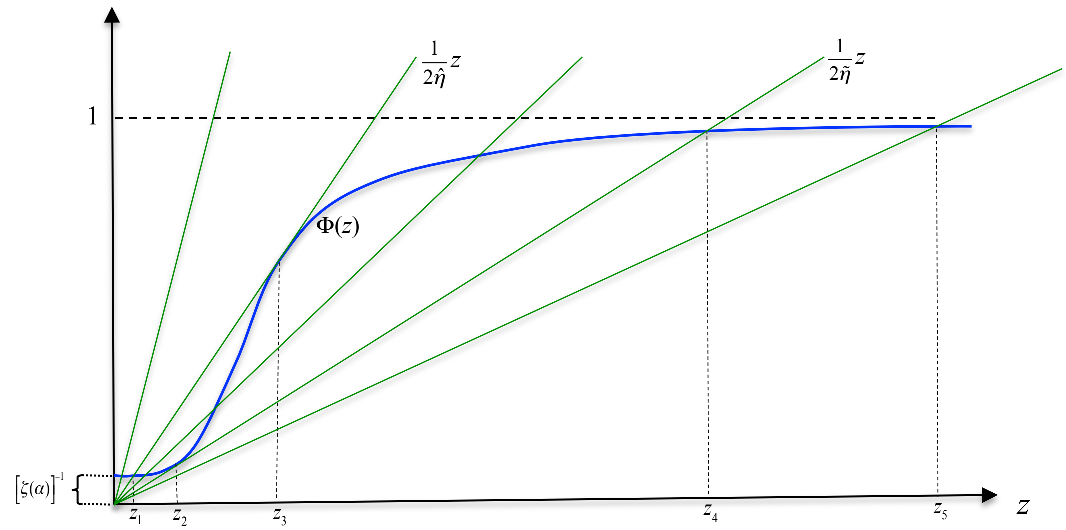


Figure 5: The effect of changes in η on the equilibrium degree under moderately high cohesion (i.e. $1 < \alpha < \bar{\alpha}$ for some relatively low $\bar{\alpha}$). See the text for an explanation.

Proposition 5 *There exists some $\bar{\alpha}$ such that if $1 < \alpha < \bar{\alpha}$, the function $z^*(\cdot)$ displays an upward right-discontinuity at some $\eta = \eta_0 > 0$.*

Remark 1 So far, the parameter μ that we have identified with the *quality of institutions* has played no explicit role in our analysis. In effect, however, it does have a very significant role in the model, by contributing to shape the function $\Phi(\cdot)$ that determines conditional linking probabilities. For example, our assumption A3 – which posits that those probabilities remain uniformly positive as the population size n becomes arbitrarily large – has to be interpreted as reflecting some auxiliary assumptions on how μ changes with n in the limiting exercise.[13] However, for the focused purposes of the present stylized ("benchmark") approach, such auxiliary assumptions are best left unspecified. Instead, we postpone a detailed consideration of the role of institutions in our model to Section 4.3, where the analysis is carried out in a finite-population environment and hence the effect of institutions is easier to grasp.

4.3 From the benchmark set-up to the general model

As explained above, the stylized approach undertaken in Subsection 4.2 is not well suited to the study of issues that are best addressed in a finite-population scenario. One of these issues, the role of institutions in

[13] By way of illustration, suppose α is very low so that, heuristically, we can identify the linking probability $\Phi(\cdot)$ to the probability that two potential partners are selected to lie in the same component of the network *and* their social distance is no higher than μ. In order for such a probability to be positive, we should have some component of significant fractional size (a requirement associated to the connectivity of the network, as captured by z) and the network distance within such component of order no higher than μ. Clearly, a way to guarantee the latter requirement is to make sure that the diameter of that component be no higher than μ. Suppose now that, as maintained throughout, the social network can be conceived as a random network. It then follows from standard theory (see the references mentioned in Footnote 11) that the diameter of the largest (so-called *giant*) component grows "slowly" with n – specifically, logarithmically. The assumption, therefore, that μ grows in n at least as fast as that would be sufficient for our purposes, i.e. would induce a positive linking probability if z is large enough.

the model, was already referred to in Remark 1. Another important phenomenon which is best studied in this scenario is the effect of geographical cohesion on the local saturation of (non-redundant) opportunities. More generally, an important limitation of the benchmark set-up is that, given the abstract nature of its postulated assumptions (A1 to A4), it permits the establishment of qualitative conclusions but not an exhaustive comparative analysis of parameter changes. To address these manifold concerns is the main objective of this subsection.

The approach pursued in this section, however, builds crucially upon the conceptual and methodological insights obtained from our preceding analysis in Subsection 4.2. In particular, we shall continue to rely on the steady-state condition (2) – and its reformulation in (3) – to determine the values for average connectivity around which the system gravitates over time. Just as before as well, this will allow us to single out the stable steady states and the transitions between them that are induced by small changes in the parameters of the model. But of course, in order to be able do so, we must first develop a way of deriving the function $\Phi(z)$ that gives the effective linking probabilities associated to different values of z. Since we now need a precise and global determination of this function for any parameter configuration and any population size, we shall do it numerically rather than analytically. Thus, in contrast with our former approach, we shall rely on numerical methods to arrive at an accurate estimate of the function Φ. We start, therefore, by describing in some detail the numerical procedure used to determine the function $\Phi(\cdot)$ – for further details, the reader is referred to Appendix C.

Algorithm P: Numerical determination of $\Phi(z)$

Given any parameter configuration and a particular value of z, the value $\Phi(z)$ is determined through the following two phases.

1. First the process is simulated starting from an empty network but putting the volatility component on hold – that is, avoiding the destruction of links. This volatility-free phase is maintained until the average degree z in question is reached.
2. Thereafter, a second phase is implemented where random link destruction is brought in so that, at each point in time, the average degree remains always equal to z. In practice, this amounts to imposing that, at each instant in which a new link is created, another link is subsequently chosen *at random* to be eliminated. (Note that, by choosing the link to be removed in an unbiased manner, the topology of the networks so constructed coincides with that which would prevail if the resulting configuration were a genuine steady state.) As the simulation proceeds in the manner described during this second phase of the procedure, the fraction of times that a link is actually created between meeting partners is recorded. When this frequency stabilizes, the corresponding value is identified with $\Phi(z)$.

Given the function $\Phi(\cdot)$ computed through Algorithm P, the value $\eta\,\phi(z)$ induced for each z acts, just as in the benchmark model, as a key anchor of reference. For, in effect, it specifies the "notional" rate of link creation that would ensue (normalized by population size) **if** such average degree z were stationary. Indeed, when the overall rate of project destruction $\lambda\frac{z}{2} = \frac{z}{2}$ equals $\eta\,\phi(z)$, the former conditional "if" applies and thus a steady state actually obtains.

Diagrammatically, the situation can be depicted as for the benchmark model, i.e. as a point of intersection between the function $\Phi(\cdot)$ and a ray of slope equal to $1/(2\eta)$. Figure 6 includes four different panels where such intersections are depicted for a fixed ratio $1/(2\eta)$ and alternative values of α and μ, the corresponding functions $\Phi(\cdot)$ depicted in each case being determined through Algorithm P.

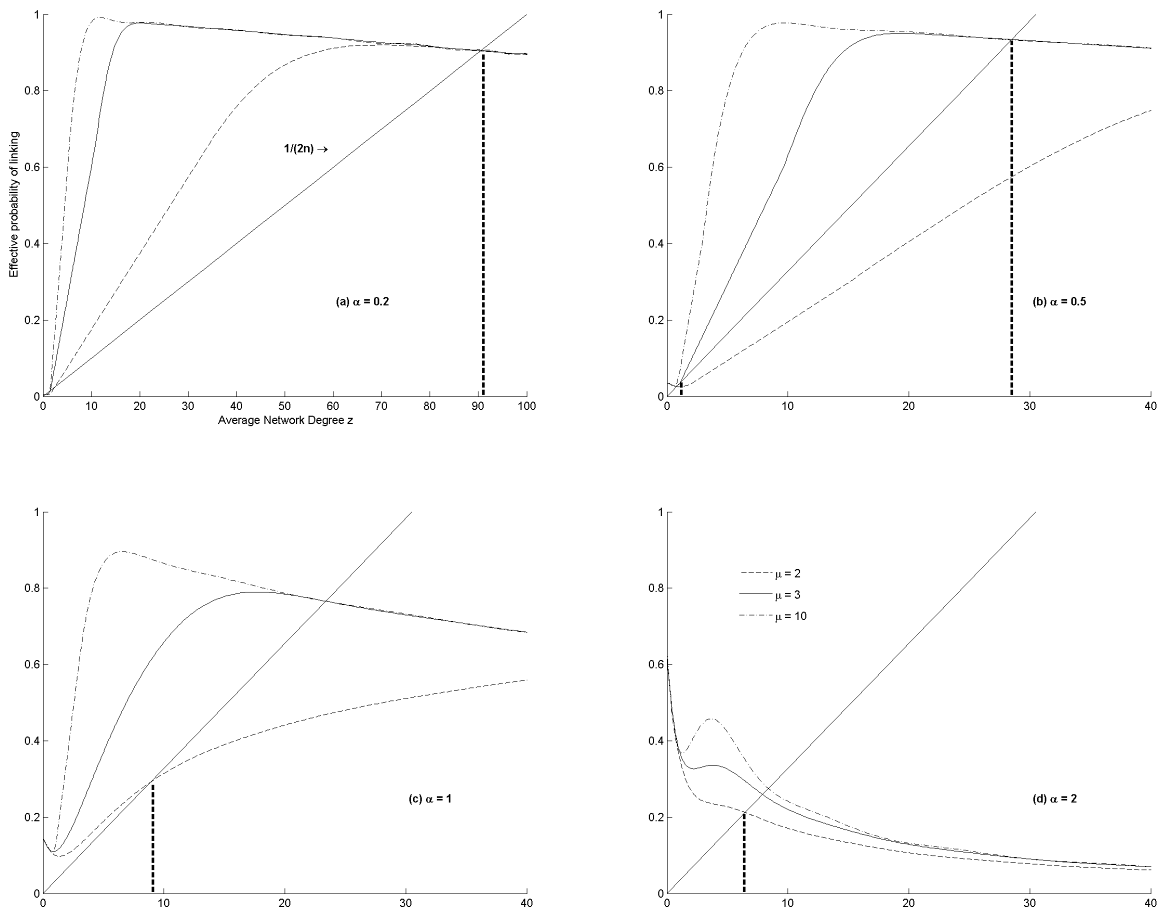


Figure 6: Graphical representation of the equilibrium condition $\Phi(z) = \frac{1}{2\eta} z$ in the general framework. The diagrams trace the steady states as points of intersection between a fixed ray of slope $1/(2\,\eta)$ and the function $\Phi(\cdot)$ computed for different institutions μ (within each panel) and cohesions α (across panels).

As a quick and informal advance of the systematic analysis that is undertaken below, note that Figure 6 shows behavior that is fully in line with the insights and general intuition shaped by our preceding analysis. First, we see that the transition towards a highly connected network is abrupt and large for low values of α, but gradual and significantly more limited overall for high values of this parameter. To focus ideas, consider specifically the panel for $\alpha = 0.5$ and the case with $\mu = 3$. There, at the value of η associated to the ray being drawn ($\eta \simeq 15$), the system is at a point of a discontinuous transition. This situation contrasts with that displayed in the panels for larger α – see e.g. the case $\alpha = 2$ – in which changes in η (affecting the slope of the ray) would trace a continuous change in the equilibrium values.

It is also worth emphasizing that some of the details apparent in Figure 6 point to significant differences with the benchmark model. For example, $\Phi(z)$ does *not* uniformly vanish to zero below a certain positive threshold $\hat{z}$ for $\alpha \leq 1$. This, of course, is simply a consequence of the fact that this condition must be expected to hold only in the limit as the population size $z \to \infty$. For finite populations, that is, significant deviations from it must be expected, which is a point that underscores the benefits of developing an approach that extends our original benchmark model to account for finite-population effects.

The question may still be raised as to whether such numerical approach indeed represents a valid way to solve for the steady-state values associated to different parameter configurations. To confirm its validity, we have conducted a systematic comparison of the theoretical predictions induced by this approach and corresponding simulation results. Rather than providing here an exhaustive account of this exercise, we illustrate matters in Figure 7 by focusing on the two scenarios (with low and a high cohesion) for which we conducted the simulations reported in Subsection 4.1 – for clarity, we report only the results that pertain the steady-state average degree. We find that, for both scenarios, there is a precise correspondence – not just qualitative but also *quantitative* – between the theoretical predictions and the simulation results. This illustrates that, indeed, we can regard our approach as a suitable way of analyzing the model, also for finite-population contexts.

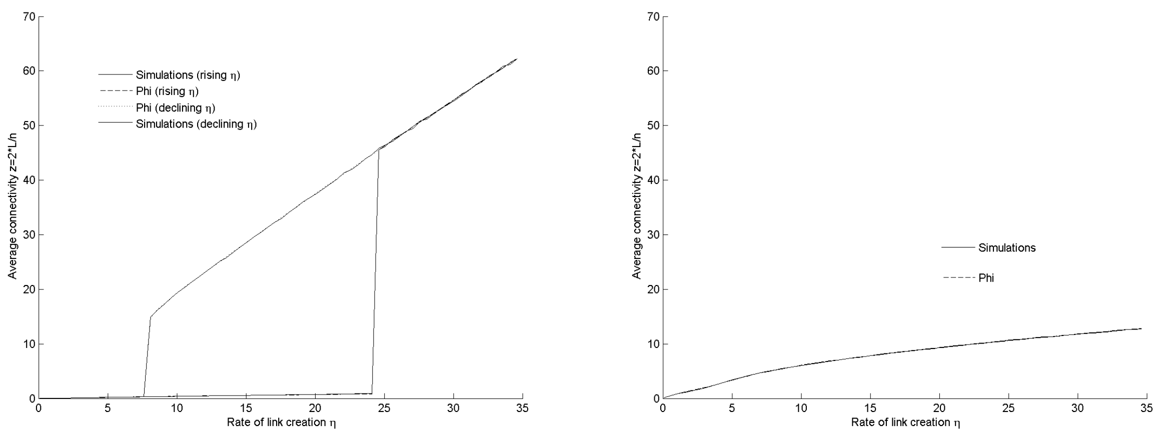


Figure 7: Comparison between the theoretical predictions and the numerical simulations on the steady-state average degree for each of the two scenarios ($\alpha = 0.5\,,\, 2$) considered in Subsection 4.1.

In the remaining part of this section, our analysis focuses on three issues. Firstly, we revisit the relationship between cohesion and the rise of globalization. Secondly, we study the effect of institutions on long-run connectivity. Finally, we address the issue of how the optimal level of cohesion depends on the remaining parameters of the model. [14]

[14]Given that our approach to determining the function $\Phi(\cdot)$ is numerical, we need to discretize the parameters η and α (institutions μ is already defined to be discrete). For η we choose a grid of unit step, i.e. we consider the set $\Psi_\eta = \{1, 2, 3, ...\}$, while for α the grid step is chosen equal to 0.05, thus the grid is $\Psi_\alpha = \{0.05, 0.1, 0.15, ...\}$. As population size, our results below are reported for $n = 1000$. We have conducted, however, robustness checks to confirm that the gist of our conclusions is unaffected by the use of finer grids or larger population sizes.

4.3.1 Geographical cohesion and the transition to globalization

Within the limit setup given by the benchmark model, the transition to globalization from a sparse network is only possible if the cohesion parameter $\alpha > 1$ (cf. Proposition 2). This, however, is strictly only true in the limit case of an infinite population, as well illustrated by the simulations reported in Subsection 4.1. However, in general, for a large but finite population, one would expect that the transition to globalization becomes progressively harder (in terms of the innovation rate η required) as α approaches, and falls below, unity. And, in line with the main insights captured by Propositions 4 and 5, it would be expected as well that, once globalization has taken place, its implications are sharper and more substantial (e.g. on network connectivity) the lower is geographical cohesion.

Our present approach permits the exploration of the former conjectures in a systematic manner across all regions of the parameter space. The results of this endeavor are summarized in Figure 8 for a representative range of the parameters α and μ. In the different diagrams included there, we trace the effect of η on the *lowest* average network degree that can be supported at a steady-state configuration of the system. In line with our discussion in Subsection 4.2, this outcome is interpreted as the long-run connectivity attainable when the social network must be "built up from scratch," i.e. from a very sparse network.

Diagrams (a)-(b) in Figure 8 show that a discontinuous and powerful transition occurs for values of α that are well below unity, while its Diagrams (e) and (f) show that the transition is gradual and much less effective if α is significantly higher than one. This contrasting behavior occurs for a significant span of values of μ, the quality of institutions. Instead, for a middle range of the cohesion parameter, which includes $\alpha = 1$ and other close values, the nature of the transition is less extreme and, in fact, changes from being discontinuous to continuous as μ increases. This is well in line with the conjectures suggested above.

4.3.2 The role of institutions

Now we turn to exploring the impact of institutions on the rise of globalization. The situation is described in Figure 9, where we show the effect of institutions on the lowest network connectivity supportable at a steady state for a representative range of the remaining parameters, α and η. The primary conclusion following from in Figure 9 is that, provided η is large enough, there is a relatively low value of μ at which institutions saturate their ability to increase the long-run connectivity of the network. That is, all what is required to achieve the full potential of institutions is that μ be above a certain threshold that is low relative to population size.

The intuition for such an "institutional saturation" is based on fact that, as explained, the social network prevailing in a steady state can be regarded as a random network. Hence we may invoke the standard theory of random networks to argue that, for sufficiently connected networks, *almost all* nodes must lie (with very high probability) at a network distance that coincides with the network diameter, i.e. they lie at *maximum* distance (cf. Footnotes 8 and 11 for useful references). This implies that, at the point where globalization occurs, μ can be no lower than the network diameter and hence further improvements in institutions should entail no further formation of links.

On the other hand, the reason why such a diameter-induced threshold should be low relative to population

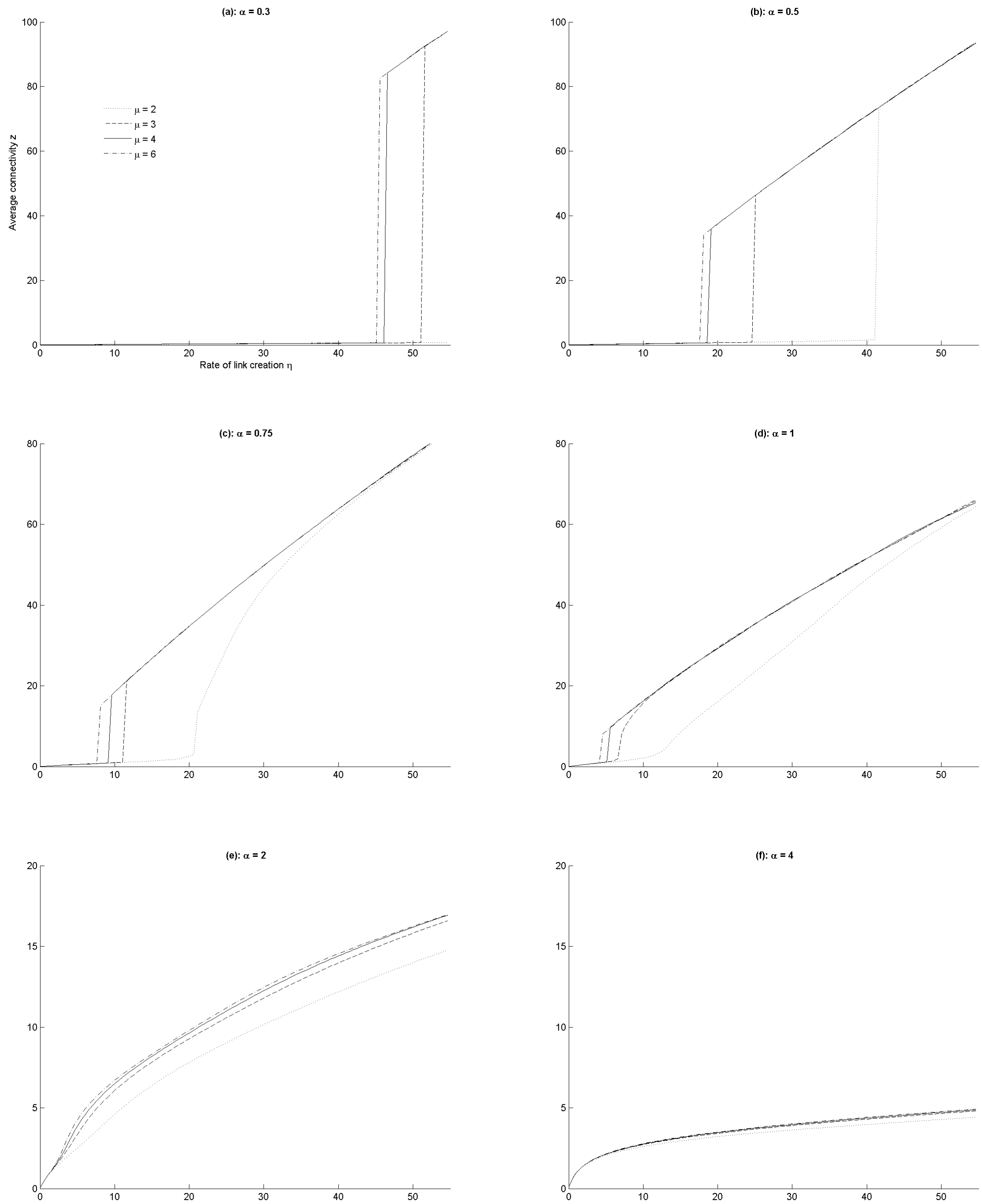


Figure 8: Numerical solution for the lowest average network degree that can be supported at a steady state, as the innovation rate η rises, for different given values of geographical cohesion α and institutions μ (with $\lambda = 1$, $n = 1000$).

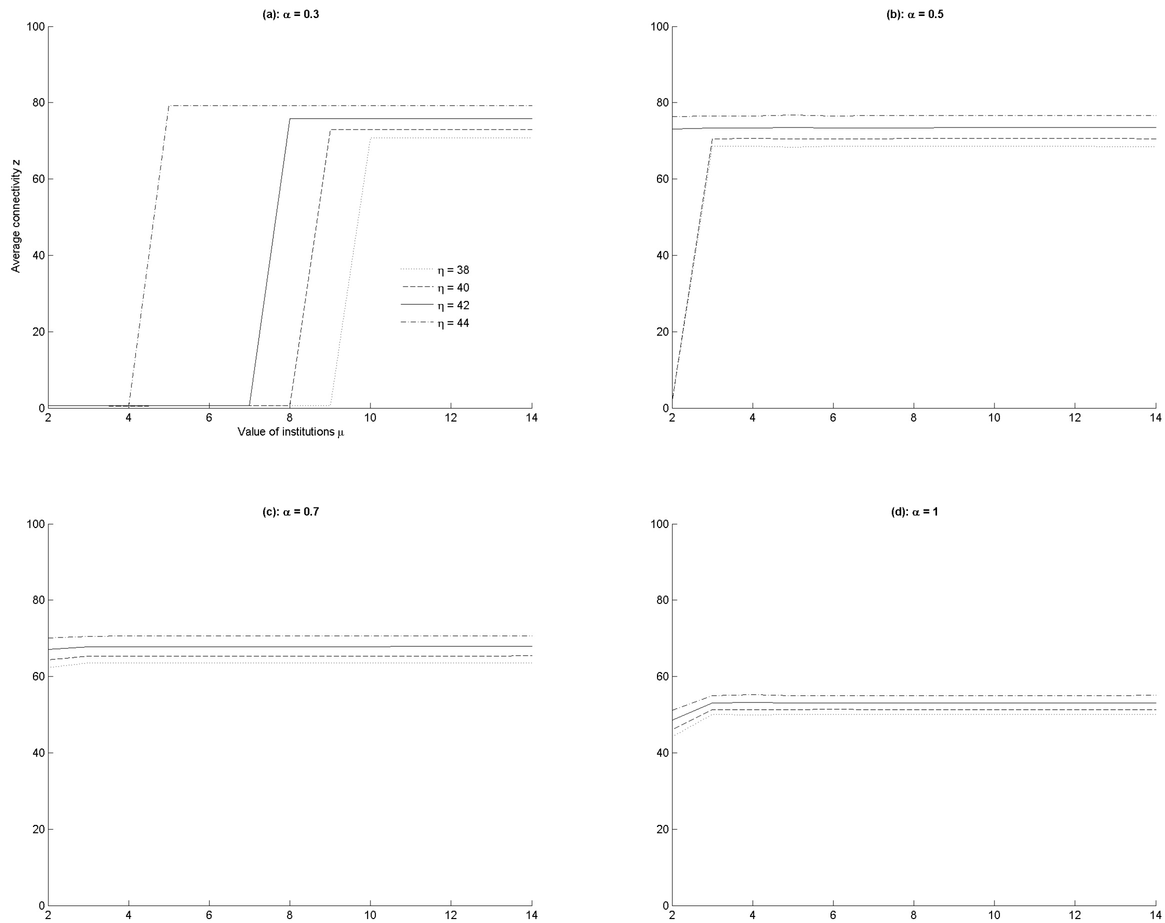


Figure 9: Numerical solution for the lowest average network degree that can be supported at a steady state, as institutions improve, for different given values of the innovation rate η and geographical cohesion α (with $\lambda = 1$, $n = 1000$).

size is less apparent. In part, it may be understood as a reflection of another prominent property of random networks that was already discussed in Remark 1: the diameter of random networks grows very slowly (i.e. scales logarithmically) with population size. This suggests that, indeed, the minimum value of institutions required to *sustain* a globalized state must be small relative to population size. In addition to this, Figure 9 shows something stronger: the value of institutions at which a globalized state *arises* also grows very slowly with population size. This fact is not just a manifestation of the local dynamic effects supporting a stable steady-state but the outcome of global dynamic forces that trigger large-scale reconfigurations of the system. That these forces become activated “once and for all” when institutions reach relative low thresholds appears quite remarkable.

Finally, let us highlight another interesting observation gathered from the above diagram. Figure 9 shows that, *for low levels of cohesion*, the effect of institutions on the average degree is step like, i.e. not only does it have no effect above the required threshold but the same happens for any improvement that keeps the resulting institutions below it. Again, the intuition for this behavior is based on aforementioned considerations. When cohesion is low, any significant flow of link creation must rely on the existence of a global network. Consequently, if institutions are not high enough to allow for such a network to arise, the rate of link creation is vanishing small and the average degree induced at a steady must remain vanishing small at all such levels of μ. Instead, at the threshold where an institutional improvement does lead to globalization, we have already explained that μ cannot lie below the network diameter. Such a threshold, therefore, also represents the saturation point beyond which any further increase is ineffectual. A combination of the two previous points leads to the step-like functions observed in Figure 9 for low levels of cohesion.

4.3.3 The optimal level of geographical cohesion

Throughout our discussion of the model, we have stressed the two contradictory implications of geographical cohesion: it endows the linking process with useful structure but, on the other hand, exacerbates the problem of congestion/saturation of *local* linking opportunities. This is why we expect that, in general, there must be an optimal resolution of this trade-off at some intermediate level. To be precise, define the Optimal Geographic Cohesion (OGC) as the optimal value $\alpha = \alpha^*$ that maximizes the long-run average network degree when the process starts from the empty network. It is natural to conjecture that such OGC should decrease with the quality of the environment, i.e. as institutions get better (higher μ) or innovation proceeds faster (higher η). This is precisely the gist of the message conveyed by Figure 10.

The intuition for the clear-cut behavior of the OGC displayed in Figure 10 should be quite clear by now. In general, geographical cohesion can only be useful to the extent that it helps the economy build and sustain a dense social network in the long run. Except for this key consideration, therefore, the lower is the geographical cohesion of the economy the better. Heuristically, one may think of the OGC as the lowest value of α that allows the build-up of a dense social network. Thus, since such a build-up becomes easier to perform the better the underlying environmental conditions, that there should be a negative dependence of OGC on η and μ is full in line with intuition.

5 Summary and conclusions

The paper has proposed a "spatial" theoretical framework to study the relationship between globalization and economic performance. The main feature of the model is that *connections breed connections*, for it is the prevailing social network that supports the materialization of new linking opportunities. In the end, it is the fast exploitation of those opportunities that offsets the persistent process of link decay and hence may sustain a steady state with a dense network of connections.

But in order for such a process to unfold, the social network must become global, i.e. network distances must be short and links must span long geographical distances. Otherwise, only local opportunities become available and thus economic performance is sharply limited by local saturation. To understand the way in which this phenomenon of globalization comes about has been the primary concern of the paper. We have seen, for example, that it may occur quite abruptly and that, as it unfolds, the interplay of "geography" and

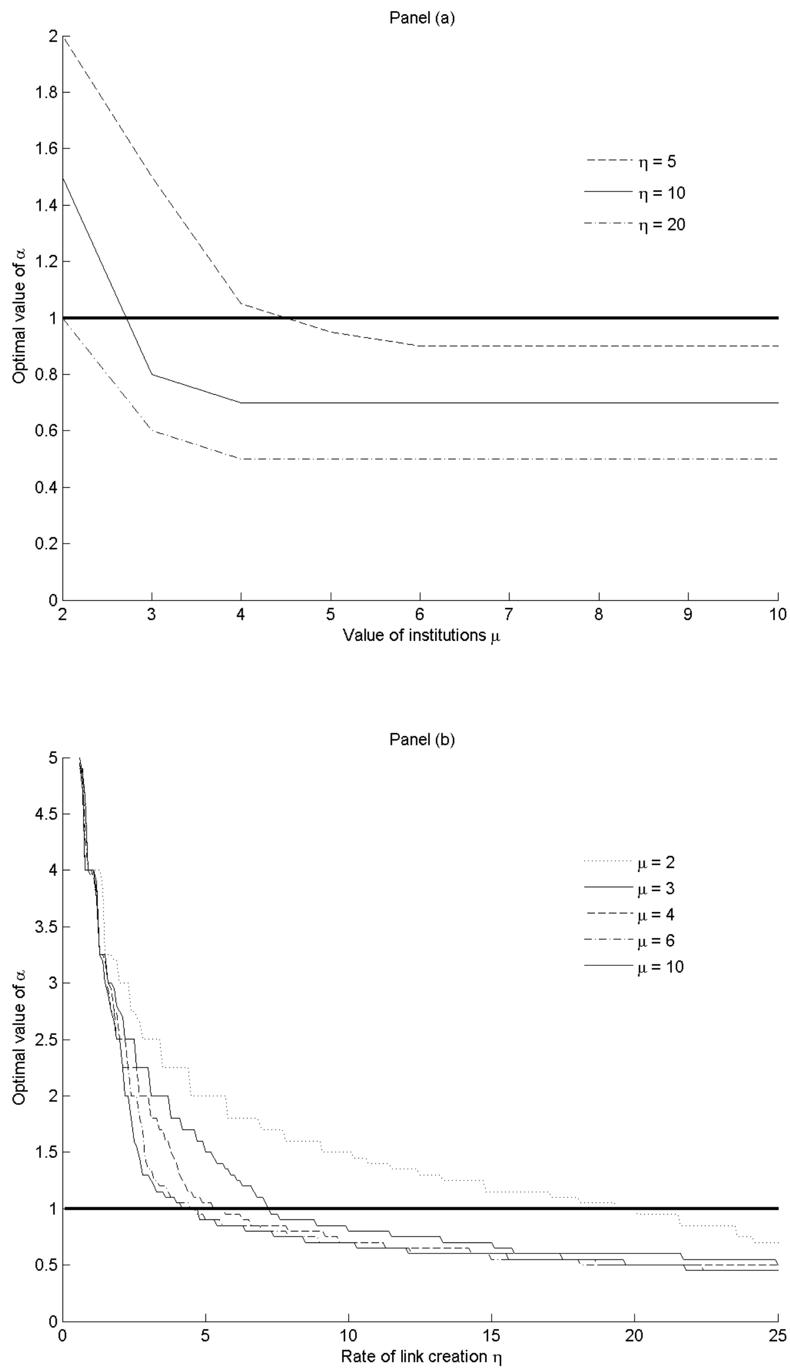


Figure 10: Optimal level of geographical cohesion (OGC), as a function of institutions μ, given different values of η (Panel (a)) and as a function of the the innovation rate η, given different values of η (Panel (b)).

the evolving social network may be quite subtle.

This paper is to be viewed as a first step in what we hope might be a multifaceted research program. From a theoretical viewpoint, an obvious task for future research is to enrich the microeconomic/strategic foundations of the model. This will require, in particular, to describe in more detail how information flows through the social network, and the way in which agents' incentives respond to it. Another important extension of the model would be to allow for significant agent heterogeneity (the basis for so much economic interaction), which should probably be tailored to the underlying geographic space, as in the work of Dixit (2003) – cf. Section 2.

The theoretical framework also opens the room for the discussion of policy. By way of illustration, recall that, if geographical cohesion is not too strong, the model yields significant equilibrium multiplicity inducing a substantial gap (say, in network connectivity) between the different equilibria. Such a multiplicity, in turn, brings in some hysteresis in how the model reacts to small parameter changes. Specifically, slight modifications in some of the parameters (e.g. a small stimulus to the innovation rate) may trigger a major shift in globalization that, once attained, would remain in place even if the original parametric conditions were restored. Conceived as a *short-run* policy measure, the effectiveness of any such *temporary* involvement derives from its ability to trigger the chain of self-feeding effects that are inherent to network phenomena. In this indeed happens, even a limited intervention may attain major changes in the system that are also robust, i.e. "locally irreversible" and hence likely to be persistent.

Finally, another important route to pursue in the future is of an empirical nature. As discussed in Section 2, there is a substantial empirical literature on the phenomenon of globalization but a dearth of complementary theoretical work supporting these efforts. The present paper hopes to contribute to closing the gap, by suggesting what variables to measure and what predictions to test. Specifically, our model highlights network interaction measures that stress the importance of both direct and indirect existing connections in the process of creating and supporting new ones. As briefly outlined, Duernecker *et al.* (2010) adopt this perspective in constructing network measures of economic interaction among countries that are in turn used to explain inter-country differences of performance across time. Conceivably, a similar approach could be used to study many other instances of economic interaction where the network dimension seems important, e.g. the internal network (formal and informal) among workers and managers in a firm (see Burt (1992)), or the network among firms collaborating in R&D projects within a certain industry (see Powell *et al.* (2005)).

Appendix

Appendix A
Microfoundation of the model: a game-theoretic approach

In an abstract manner, any prevailing link connecting two agents was conceived in Section 3 as an ongoing project that is run continuously until it becomes obsolete and disappears. As explained in Subsection 3.2, the general rules L1-L2 by which new links are formed admit diverse interpretations. Here we propose a game-theoretic framework describing in detail how projects are set-up and operated, which provides a possible microfoundation for those rules.

Consider any two agents, i and j, who are facing an opportunity to start one such project. This project amounts to initiating a bilateral interaction that is modelled as a dynamic game (with an uncertain end) and consists of two different phases.

1. **Setup Stage** First, at the start of the interaction, the agents have to finance some fixed cost $2K$ to setup the project. Only if they manage to cover this cost the project can start to operate. To tackle the problem, they play what essentially amounts to a one-shot Prisoner's Dilemma. Each of them has to decide, independently, whether to *Cooperate* (C) or *Defect* (D). If both choose C, they share the cost equally (i.e. each covers K) and the project starts. Instead, if only one cooperates, the cooperator covers the total cost $2K$ while the defector pays nothing, and the project can subsequently start to operate as well. Finally, if both defect, the setup cost is not covered, the project fails to start, and the economic opportunity is irreversibly lost.

2. **Operating Phase** Provided that the setup cost has been covered in the first stage, the project immediately starts to run over (continuous) time until the point when it becomes "obsolete" and vanishes (recall Subsection 3.2.2). During its lifetime, the project delivers a return *flow* that depends on the effort exerted by the two agents involved. At each instant of time, their efforts can be high (H) or low (L). While an agent who exerts high effort incurs an (additive) cost of b, low effort is costless. If both agents choose H, then each obtains a *gross* payoff flow of $1 + b$ and hence a *net* payoff of 1. Otherwise (i.e. if at least one agent exerts low effort), the project generates a gross payoff flow of W (< 1) for each agent, so that the net payoff flow in that case is W for the agent(s) choosing L and $W - b$ for the agent (if any) who chooses H. To sum up, we can describe the situation within the operating phase as the repeated play of a simultaneous stage game displaying the following payoff table:

i \ j	L	H
L	W, W	$W, W - b$
H	$W - b, W$	$1, 1$

with a stopping probability rate of $\lambda = 1$.

To complement the above description of the situation, assume that the fixed cost K incurred in the Setup Stage is dependent on the geodistance of the two agents involved, i and j. Specifically, let us make the simplifying assumption that this cost solely depends on whether they are neighbors (i.e. $d(i, j) = 1$) or not:[15] In the former case, they incur a relatively low cost K_0 (say, because their geographic proximity allows them to reduce the setup cost), while otherwise the setup cost incurred is $K_1 > K_0$.

Now we introduce two different strategic scenarios in which one can study the intertemporal interaction described in Phases 1 and 2:

Bilaterally Independent (BI) Scenario: In this case, each network link defines a bilateral repeated game that is strategically independent of what occurs anywhere else in the economy. Hence overall population play is merely a juxtaposition of the equilibrium behavior displayed by each of the connected couples in their respective bilateral games.

Population-Embedded (PE) Scenario: This corresponds to a situation where the behavior in the bilateral interaction of two players can be made dependent on (and also affect) the behavior displayed in some other third-party interaction undertaken by these players. Under these conditions, therefore, the whole population must be seen as playing a single common game.

Next, we compare the implications of each of the two scenarios, BI and PE. In both cases, the central issue is whether repeated interaction (either for a particular couple of partners or across different couples) can provide agents with *individual* incentives to cover the setup cost involved in the creation of new links. To render the problem interesting, we assume that it can never be in the interest of a single agent to pay the full amount $2K$ required to start the operation of a project, but instead it is worthwhile to cover such a setup cost if shared equally. To be precise, suppose that agents discount future payoffs at the common rate $r > 0$, so that the effective discount rate within a given ongoing project is $\rho \equiv r + \lambda = r + 1$. Then, the required condition is as follows:

$$K < \frac{1}{\rho} < 2K. \tag{12}$$

We assume that the previous inequalities are satisfies for any of the projects that may arise, be they between neighbors or not. This amounts to positing the following.

G1. $K_1 < \frac{1}{\rho} < 2K_0$.

The former inequality implies that, within the payoff possibilities provided by a single interaction, the only way in which both agents can be interested in starting a project is if they can anticipate that the partner will share the cost (at least with some probability). Given the nature of the the setup game, however, such a state of affairs will not materialize unless the equilibrium in place punishes any deviant from cooperation in that first stage of the relationship. Such a punishment may occur, of course, within the ensuing bilateral relationship between the two agents involved in the link. This requires that there is a continuation equilibrium in the second phase of their relationship where the induced payoff consequences can be made severe enough. Formally, it amounts to postulating the following condition:

[15] This is in line with with the assumption that $nu = 1$ and could be readily adapted if any other value of ν were to be considered.

$$\frac{W}{\rho} < \frac{1}{\rho} - K \tag{13}$$

If the above condition holds, link creation decisions may be supported bilaterally, hence rendering the problem independent of network considerations. So let us start by assuming that (13) is satisfied for geographic neighbors, i.e.

G2. $\frac{W}{\rho} < \frac{1}{\rho} - K_0$.

This assumption implies that neighbors can benefit from the lower setup cost K_0 to support their collaboration in every new project at equilibrium.

On the other hand, concerning agents who are *not* neighbors, the interesting assumption to make is that condition (13) does *not* apply. This implies that their cooperation in the Startup Phase can only be attained if their behavior is socially "embedded". The simplest case in which such social embeddedness can be effective is when

$$\frac{2W}{\rho} < \frac{2}{\rho} - K. \tag{14}$$

The above condition, which is is obviously weaker than (13), allows for cooperation to be supported as individually optimal at the setup stage if, after a deviation from it, not only the partner but also an additional agent punishes the deviator. We shall make, therefore, the following assumption:

G3. $\frac{2}{\rho} - 2K_1 < \frac{2W}{\rho} < \frac{2}{\rho} - K_1$.

This assumption implies that, in order to have cooperation among non-neighboring agents, a minimal amount of "social embeddedness" of the relationship is sufficient.

Under the parameter restrictions embodied by G1, G2, and G3 (which can be easily seen to be compatible),[16] we construct an equilibrium of the repeated game played by the population over time with the following characteristics:

(i) Agents are forward looking in every respect, except that they do not envisage how their behavior at any given t may have an effect on the future evolution of the network, $g(\tau)$ for $\tau > t$. That is, they abstract from the process of network formation and behave as if the current network were to remain in place forever.

(ii) Information on the behavior of some agent i at the setup stage of any new project starting at t spreads instantaneously through the network component in $g(t)$ to which i belongs.

Building upon (i)-(ii), we shall consider equilibria of the population game where every player $i \in N$ chooses a particular form of "trigger strategy." To define it precisely, we need to introduce some notation. First, given some $\mu \in \mathbb{N}$, a social network g, and any pair of agents, i and j, such that $ij \notin g$ denote by $\mathcal{N}_{ij}^{\mu}(g) \equiv \{k \neq j : \; ik \in g \wedge \delta_g(j,k) \leq \mu - 1\}$ as the set of those neighbors of i who are no farther away than

[16]Note that if one makes $K_1 = 2K_0 = \frac{1}{\rho}$ and $W = \frac{1}{2}$, the inequalities in G1-G3 are satisfied with equality. Having observed this, it is immediate to see that small perturbations from those values can be implemented that lead to the desired inequalities.

$\mu - 1$ from j (who is not a neighbor of i). Of course, in general, this set could be empty. But if it is not, a particular agent in this set, $\varsigma_{ij}^{\mu}(g)$, is assumed to be selected in some arbitrary but pre-established manner. As explained below, this agent is the one who is in charge of punishing i if he defects upon j in the setting-up of the link ij.

Relying on the previous notation, we posit that every agent i in the population relies on a strategy that prescribes the following (contingent) behavior:

- Consider first the decisions to be taken at any t concerning any given project that is already in the operating phase between i and some other agent j. Then, agent i is taken to choose L (low effort) if, and only if, any of the following three contingencies apply:
  - either i or his partner j defected in the set-up stage of the project;
  - either i or/and his partner j chose L in some of the previous operating stages of the project;
  - a new project has started at t that involves i (respectively j) in which this agent has *unilaterally* defected in the setup stage upon some third player k and $\varsigma_{ik}^{\mu}(g(t)) = j$ (respectively, $\varsigma_{jk}^{\mu}(g(t)) = i$), where $g(t)$ is interpreted as the network prevailing at t prior to the formation of any link that is created at that point in time.
- Consider now the decision to be taken at the setup stage of a project that becomes available at t between i and some other j. Then, agent i chooses C (cooperation) in this stage if, and only if, either i and j are geographic neighbors or $\delta_{g(t)}(i,j) \leq \mu$.

It is straightforward to check that, if agents' perceptions and information are as described in (i)-(ii) and stage payoffs satisfy assumptions G1-G3, the strategy described above defines a symmetric Nash equilibrium of the population game. This equilibrium reflects a cooperative *social norm* where agents cooperate not only with neighbors but also with all those who are socially close (as specified by μ). On the other hand, they are also willing to punish deviators on third parties provided these are close enough (the upper bound in this case being $\mu - 1$). The implication of all this is that, since agents will anticipate this behavior at the time of forming new links, they will be ready to form those links that connect to either geographic neighbors or to all those agents who are not more than μ links apart. This is precisely the link-formation rule that was postulated in Subsection 3.2.1.

In our general description of the model, we have identified the parameter μ with the (quality of the) institutions prevailing in the economy. In abstract terms, it captures the extent to which agents can rely upon the social network to support the formation of new links/projects. As suggested, it could reflect how fast information travels along the network, or how useful it is in providing access to new opportunities. In the context of our present strategic framework, μ can be interpreted along the lines of what Karlan *et al.* (2009) label the "circle of trust," which is a notion that has been highlighted by a number of prominent scholars studying the effect of institutions on economic performance.[17] The essential idea here is that μ

[17] An early instance can be found in the celebrated study of Southern Italy by Banfield (1958). There he argued that the persistent backwardness of this region was largely due to a pervasive *amoral familism*, i.e. the *exclusive* concern for the well-being of the closely related, as opposed to that of the community at large. More recently, Platteau (2000) has elaborated at length on the idea, stressing the importance for economic development of the dichotomy between *generalized morality* (moral sentiments applied to abstract people) and *limited-group morality* (which is restricted to a concrete set of people with whom one shares a sense of belonging).

measures how strongly agents "internalize" the social norm of cooperation. If the support of this norm – conceived as the quality of institutions – is weak, agents should only be prepared to punish others if they have defected upon individuals who are socially very close. Instead, if institutions are strong, any violation of the social norm that agents are acquainted with – even if it involves agents who are socially distant – is readily punished.

Appendix B
Proofs

Proof of Proposition 1

First, in order to establish (7), note that, upon receiving a link-formation opportunity, a necessary condition for any agent i to be able to establish a new link at t with some other agent j who has been selected as given in (1) is that either j belongs to the same network component of i or/and both nodes are (geographical) neighbors. Given that there are $M_i - 1$ other nodes in i's component and every agent has two geographic neighbors, the desired upper bound follows from the fact that the maximum probability with which any agent j can be chosen as a potential partner of i is $[2\,\zeta(\alpha, n)]^{-1}$.

On the other hand, to establish (8), we simply need to recall that an isolated node i will *always* form a new link with either of its two geographical neighbors if these are selected as potential partners. Since the geographic distance to each of them is normalized to 1, the probability that either of them is selected is simply 2 $[2\,\zeta(\alpha, n)]^{-1}$. This readily leads to the desired lower bound. ■

Proof of Proposition 2

Given α, define by $F(z) \equiv \eta\,\Phi(z;\,\alpha) - \frac{z}{2}$ the one-dimensional vector field defining the dynamical system (10). By A1, when $\alpha \leq 1$, the derivative $F'(z)$ of this function at any $z < \hat{z}$ is given by:

$$F'(z) = -\frac{1}{2} < 0 \tag{15}$$

which obviously induces the desired conclusion. ■

Proof of Proposition 3

In view of A2, when $\alpha > 1$, we have that $\Phi(0, \alpha) > 0$ provided $\eta > 0$. This implies that $z = 0$ is not a steady state of the system (10), hence *a fortiori* not asymptotically stable.

Let now denote by z^* the lowest value of z such that $\eta\,\Phi(z;\,\alpha) = z/2$. From A3, such z^* is well-defined and strictly positive. It then follows that, for some sufficiently small initial $z(0)$, $\varpi\,(z(0), t) \to z^*$, as claimed. ■

Proof of Proposition 4

Consider any arbitrarily large value $\bar{z}$. Given α, let $\chi \equiv \min\{\Phi(z, \alpha) : z \leq \bar{z}\}$. From A4, $\chi > 0$. Choose

$\bar{\eta}$ such that $\bar{\eta}\,\chi > \bar{z}/2$. Then, clearly, if $\eta \geq \bar{\eta}$, we have:

$$\forall z \leq \bar{z}, \quad \eta\,\Phi(z,\alpha) > z/2, \tag{16}$$

which implies that $z^*(\eta) \geq \bar{z}$, where $z^*(\eta)$ is the unique steady introduced in Proposition 3. This completes the proof. ∎

Proof of Proposition 5

Consider any given $\hat{\alpha} > 1$. By A4, there exists some corresponding z_1 and z_2 ($z_1 < z_2$) and $\zeta > 0$ such that

$$\forall \alpha \leq \hat{\alpha},\ \forall z \in (z_1, z_2), \quad \Phi(z,\alpha) > \zeta. \tag{17}$$

This then implies that we can choose some $\hat{\eta}$ such that, if $\eta \geq \hat{\eta}$, then

$$\forall \alpha \leq \hat{\alpha},\ \forall z \in (z_1, z_2), \quad \eta\,\Phi(z,\alpha) > \frac{1}{2}z.$$

Fix now z_1 and z_2 as above as well as some $\alpha \leq \hat{\alpha}$. For any given $\beta > 0$, define the ray of slope β as the locus of points $\mathcal{R}_\beta \equiv \{(z,y) \in \mathbb{R}^2 : y = \beta\, z,\ z \geq 0\}$. We shall say that a ray $\mathcal{R}_\beta$ supports Φ on $[0, z_2]$ if

$$\forall z \in [0, z_2], \quad \Phi(z,\alpha) \geq \beta\, z$$

with the above expression holding with equality for some $\hat{z} \in [0, z_2]$, i.e. for such $\hat{z}$ we have:

$$\Phi(\hat{z},\alpha) = \beta\,\hat{z}. \tag{18}$$

Now note that, by A1 to A3, there must exist some $\bar{\alpha} > 1$ with $\bar{\alpha} \leq \hat{\alpha}$ such that, if $1 < \alpha \leq \bar{\alpha}$ the (unique) ray $\mathcal{R}_{\beta'}$ that supports Φ on $[0, z_2]$ has its slope β' satisfy

$$0 < \beta' \leq \frac{1}{2\hat{\eta}} \tag{19}$$

So choose $\bar{\alpha}$ so that (19) holds and let $\eta' = \dfrac{1}{2\beta'}$. For such value η' we may identify $z^*(\eta')$ with the lowest average degree $\hat{z}$ such that (18) holds. Clearly, $z^*(\eta') \leq z_1$, where z_1 is chosen as in (17).

Now consider any $\eta'' > \eta'$. Then we have:

$$\forall \alpha \leq \bar{\alpha},\ \forall z \in [0, z_2], \quad \eta''\,\Phi(z,\alpha) > \frac{1}{2}z. \tag{20}$$

But, since $\Phi(\cdot) \leq 1$, there must exist some z'' such that

$$\eta''\,\Phi(z'',\alpha) = \frac{1}{2}z''. \tag{21}$$

As before, we identify $z^*(\eta'')$ with the lowest value of z'' for which (21) holds. Then it follows from (20) that $z^*(\eta'') \geq z_2$. This establishes that the function $z^*(\cdot)$ displays an upward right-discontinuity at η' and completes the proof. ∎

Appendix C
Simulation algorithm

Here we describe in detail the algorithm that is used to conduct the simulations in Subsection 4.1. The adaptation of it that is used to compute numerically the function $\Phi(\cdot)$ in Subsection 4.3 (which is labelled as "Algorithm P") is based on the same procedure.[18]

The algorithm proceeds in two successive steps, which are repeated until certain termination criteria are met. The first step selects and implements a particular adjustment event (which can be either an innovation draw or a link destruction) and the second step checks whether or not the system has reached a steady state.

As mentioned, we normalize the rate of link destruction to $\lambda = 1$. Thus the free parameters of the model are η, α, and μ. The state of the network at any point in the process is characterized by the $n \times n$ dimensional adjacency matrix A. A typical element of it, denoted by $a(i, j)$, takes the value of 1 if there exists an active link between the nodes i and j, and it is 0 otherwise. L denotes the total number of active links in the network. By construction, L has to equal twice the number of non-zero elements in the state matrix A.

We now describe systematically each of the steps the algorithm. The algorithm runs in discrete steps but its formulation is intended to reflect adjustments that are undertaken in continuous time. Hence only one adjustment takes place at each step (either a new link is created or a pre-existing links is destroyed), with probabilities proportional to the corresponding rates. Some given A, characterizes the initial state, which can be either an empty network (with A containing only zeros), or some other network generated in some pre-specified manner.

- **Step I:** At the start of each simulation step, $t = 1, 2, ...$, an adjustment event is randomly selected: This event can be either an innovation draw or a link destruction. As explained, the two possibilities are mutually exclusive. The rates at which either of the two events occur are fixed and equal to λ per link and η per node. Every node in the network is equally likely to receive an innovation draw. Similarly, all existing links are equally likely to be destructed. Hence the flow of innovation draws and destroyed links are respectively given by $\eta\, n$ and $\lambda\, L$. This implies that the probability of some innovation draw to occur is $\frac{\eta n}{\eta n + \lambda L}$ whereas some link destruction occurs with the complementary probability. Depending on the outcome of this draw, the routine proceeds either to Step A.1. (innovation draw) or Step B.1. (link destruction)
  - A.1. The routine starts by selecting at random the node $i \in N$ that receives the project draw. All nodes in the network are equally likely to receive the draw, so the *conditional* selection probability for a particular node is $1/n$. Having completed the selection, the algorithm moves to A.2.
  - A.2. Another node $j \in N$ $(j \neq i)$ is selected as potential "partner" of i in carrying out the project. The probability $p_i(j)$ that any such particular j is selected satisfies $p_i(j) \propto d(i, j)^{-\alpha}$. This can be translated into an exact probability $p_i(j) = B \times d(i, j)^{-\alpha}$ by applying the scaling factor $B = \left(\sum_{j \neq i \in N} d(i, j)^{-\alpha}\right)$ derived from the fact that $\sum_{j \neq i \in N} p_i(j) = 1$. The algorithm then moves to A.3.

[18] The MATLAB code implementing the algorithm is available upon request.

– A.3. If $a(i,j) = 1$, there is already a connection in place between i and j. In that case, the innovation draw is wasted (by L1 in Subsection 3.2.1), and the algorithm proceeds to **Step II**. If, instead, $a(i,j) = 0$ the algorithm proceeds to A.4.

– A.4. In this step, the algorithm examines whether or not it is admissible (given L2 in Subsection 3.2.1) to establish the connection between i and j. First, it checks whether i and j are geographic neighbors, i.e. $d(i,j) = 1$. If this is the case, the link is created (by L2(a)) and the step ends. Otherwise, it computes the current network distance between i and j, denoted $\delta_A(i,j)$. This distance is obtained through a breadth-first search algorithm that is described in detail below. If it is found that $\delta_A(i,j) \leq \mu$ then the link ij $(= ji)$ is created (by L2(b)) and the corresponding elements in the adjacency matrix A, $a(i,j)$ and $a(j,i)$ are set equal to 1. Instead, if $\delta_A(i,j) > \mu$, the link is not created. In either case, Step A.4 is ended. At the completion of this step, the algorithm proceeds to **Step II**.

– B.1. If the event selected in **Step I** involves a link destruction, the algorithm randomly picks one of the existing links in the network and dissolves it. The state matrix is updated accordingly by setting $a(i,j)$ and $a(j,i)$ both equal to 0. All existing links in the network are equally likely to be destructed. Thus, for a specific link the probability of being selected is L^{-1}. Once the link destruction process is completed, the algorithm moves on to **Step II**.

- **Step II:** If we start with an empty network (with A containing only zeros) and let the two forces - innovation and volatility - operate, then network gradually builds up structure and gains in density. If this process is run long enough, eventually, the network attains its equilibrium. An important question in this context is, when to terminate the simulation? Or put differently, how can we find out that the system has reached a stationary state? **Step II** of the algorithm is concerned with this issue. Strictly speaking, the equilibrium of the network is characterized by the constancy of all the endogenous variables. That is, in equilibrium, the structure of the network, as measured for instance by the average connectivity, remains unchanged. However, a computational difficulty arises from the random nature of the processes involved. Link creation and destruction are the result of random processes, which imply the constancy of the endogenous variables only in expectations. In other words, each adjustment step leads to a change in the structure of the network, and consequently, the realization of each of the endogenous outcomes fluctuate around a constant value. To circumvent this difficulty, the algorithm proceeds as follows:

  – C.1. At the end of each simulation step t , the routine computes (and stores) the average connectivity prevailing in the current network as $z(t) = \frac{2 \times L(t)}{N}$.

  – C.2. Every T steps it runs an OLS regression of the $\underline{T}$ most recent values of z on a constant and a linear trend.

  – C.3. Every time the slope coefficient changes its sign from plus to minus, a counter n is increased by 1.

Steps I and **II** are repeated until the counter n exceeds the predetermined value of $\overline{n}$. For certain parameter combinations, mainly for those that imply high and globalized interaction, the transition process towards the equilibrium can be very sticky and slow. For that reason and to make sure that the algorithm does not terminate the simulation too early we set $\underline{T} = 5 \times 10^5$, $T = 10^4$ and $\underline{n} = 10$.

Breadth-first search algorithm: In Step A.4. we use a breadth-first search algorithm to determine if, starting from node i, the selected partner node j can be reached along the current network within at most μ steps. The algorithm is structured in the following step-wise fashion:

- Step $m = 1$: Construct the set of nodes which are directly connected to i, Formally, this set is given by $X_1 = \{k \in N : \delta_A(i,k) = 1\}$. Stop the search if $j \in X_1$ otherwise proceed to Step $m = 2$
- Step $m = 2, 3, ...$ For every node $k \in X_{m-1}$ construct the set $V_k = \{k' \in N \setminus \{i\} : \delta_A(k,k') = 1\}$. Let X_m be the union of these sets with all the nodes removed which are already contained in X_{m-1}. Formally: $X_m = \left\{ \bigcup_{k \in X_{m-1}} V_k \right\} \setminus X_{m-1}$. By construction, all nodes $k' \in X_m$ are located at geodesic distance m from the root i, i.e. $\delta_A(i,k') = m, \forall k' \in X_m$. Moreover, all elements in X_m are nodes that were not encountered in any of the previous $1, 2, ..., m-1$ steps. Stop the search if (a) $j \in X_m$, (b) $m = \mu$, or (c) $X_m = \emptyset$, otherwise proceed to Step $m + 1$. In Case (a), Node j has been found within distance $\mu \leq \mu$. In Case (b), continuation of the search is of no use as $\delta_A(i,j) > \mu$, in which case the creation of the link ij cannot rely on rule L2(b). Finally, in Case (c), no new nodes are encountered along the search, which implies that i and j are disconnected from each other.

The above-described search proceeds (for no more than μ steps) until Case (a), (b), or (c) occurs.

Computation of the variables of interest: In the text we report the equilibrium outcomes of four endogenous variables: the average connectivity of the network, the average geographical distance spanned by the links, the average network distance, and the effective probability of link creation. We next show how each of these are computed.

1. To compute the average connectivity of the network we simulate the equilibrium of the system for $t = 1, 2, ...\bar{t}$, with $\bar{t} = 5 \times T$, steps and take the average of $z(t)$ over all $\bar{t}$ realizations.
2. Similarly we compute the average geographical distance between connected nodes in the network as the average of $\left\{ \frac{1}{L} \sum_{ij \in g(t)} d(i,j) \right\}_{t=1}^{\bar{t}}$.
3. Formally, the average network (or geodesic) distance should be computed as, $\frac{1}{N} \sum_{i,j \in N} \delta_g(i,j)$. However, in the current context this approach is not advisable, due to the potential existence of disconnected subparts in the network. Any two nodes, i and j, which are not jointly located on such a subpart would - literally - be $\delta_g(i,j) = \infty$ steps away from one another. To account for that we randomly draw N pairs of (i,j) and compute $\delta_g(i,j)$ for each of them. If i and j happen to be disconnected we set $\delta_g(i,j)$ equal to a high number $\bar{\delta} < \infty$. The randomization also helps to economize on computational speed, since computing $\delta_g(i,j)$ for all possible pairs (i,j) would imply a substantial computational burden.
4. The effective probability of link creation is computed as the ratio of the number of innovation draws which lead to a successful link creation to the total number of innovation draws obtained in $\bar{t}$ simulation steps.

References

[1] Apostol, T. A. (1974): *Mathematical Analysis*, Addison-Wesley Pub. Co., Reading, Ma,

[2] Arribas, I., F. Pérez and E. Tortosa-Ausina (2009): "Measuring globalization of international trade: theory and evidence," *World Development* **37**, 127-145.

[3] A.T. Kearney/Foreign Policy Magazine (2006): The Globalization Index, http://www.foreignpolicy.com.

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

[5] Banfield, E.C. (1958): *The Moral Basis of a Backward Society*, New York: The Freepress.

[6] Bisin, A. and T. Verdier (2001): "The economicsof cultural transmission and the dynamicsof preferences," *Journal of Economic Theory* **97**, 289-319.

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

[8] Borensztein, E., J. de Gregorio, and J.-W. Lee (1998): "How does foreign direct investment affect economic growth?," *Journal of International Economics* **45**, 115-135.

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

[10] Centre for the Study of Globalisation and Regionalisation (2004): The CSGR Globalisation Index, http://www2.warwick.ac.uk/fac/soc/csgr/index/.

[11] Coleman, J. S. (1988): "Social Capital in the Creation of Human Capital," *American Journal of Sociology* **94**, 95-120.

[12] Coleman, J. S. (1990): *Foundations of Social Theory* , Cambridge, MA: Harvard University Press.

[13] Dollar, D. and A. Kraay (2001): "Trade, growth, and poverty," *World Bank Discussion Paper*, Washington, D.C..

[14] Dreher, A. (2006): "Does globalization affect growth?", Evidence from a new index of globalization, *Applied Economics* **38**, 1091-1110.

[15] Dreher, A., N. Gaston and P. Martens (2008): *Measuring Globalization: Gauging its Consequences*, New York: Springer .

[16] Dixit, A. (2003): "Trade Expansion and Contract Enforcement," *Journal of Political Economy* **111**, 1293-1317.

[17] Duernecker, G., M. Meyer and F. Vega-Redondo (2011): "An network-based empirical analysis of economic globalization," work in progress, University of Mannheim and European University Institute.

[18] Durrett, R. (2007): *Random Graph Dynamics*, Cambridge: Cambridge University Press.

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

[20] Fagiolo, G., J. Reyes, and S. Schiavo (2010): "The evolution of the World Trade Web," *Journal of Evolutionary Economics* **20**, 479-514.

[21] Greif, A. (1993): "Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition," *American Economic Review* **83**, 525-48.

[22] Haag, M. and R. Lagunoff (2006): "Social norms, local interaction, and neighborhood planning," *International Economic Review* **47**, 265-96.

[23] Jackson, M., T. Rodriguez-Barraquer and Xu Tan (2010): "Social capital and social quilts: network patterns of favor exchange," mimeo, Stanford University.

[24] Trajtenberg, M. and A. Jaffe (2002): *Patents, Citations and Innovations: A Window on the Knowledge Economy*, Cambridge, Mass.: MIT Press.

[25] Kali, R. and L. Reyes (2007): "The architecture of globalization: a network approach to international economic integration," *Journal of International Business Studies* **38**, 595-620.

[26] Lane, P. and G. M Milesi-Ferretti (2001): "The external wealth of nations: measures of foreign assets and liabilities for industrial and developing countries," *Journal of International Economics* **55**, 263-94.

[27] Lippert, S. and G. Spagnolo (2006): "Networks of relations, word-of-mouth communication, and social capital, SSE/EFI Working Paper in Economics and Finance no. 570. [[Check recent publication]]

[28] Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009): "Trust and Social Collateral," *Quarterly Journal of Economics* **124**, 1307-1331.

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

[30] Platteau, J. P. (2000): *Institutions, Social Norms, and Economic Development*, Amsterdam: Hardwood Academic Publishers and Routledge.

[31] 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.

[32] Sala-i Martin, X., G. Doppelhofer, and R. I. Miller (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review* **94**, 813-835.

[33] Tabellini, G. (2008): "The scope of cooperation: values and incentives," *Quarterly Journal of Economics* **123**, 905-950

[34] OECD (2005*a*): Measuring Globalisation: OECD Handbook on Economic Globalisation Indicators, June 2005.

[35] OECD (2005*b*): Measuring Globalisation: OECD Economic Globalisation Indicators, November 2005.OECD (2005*b*): Measuring Globalisation: OECD Handbook on Economic Globalisation Indicators, June 2005.

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

[37] Vega-Redondo, F. (2007): *Complex Social Networks*, Econometric Society Monograph Series, Cambridge: Cambridge University Press.