Social networks and the process of "globalization"

Georg Duernecker
University of Mannheim

Fernando Vega-Redondo
European University Institute

September 17, 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 overcoming 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. Thus only if the social network becomes globaI (heuristically, "reaches far") 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 our interesting insights is that in order for a local social network to turn global, the economy needs to display a degree of "geograhical cohesion" that is neither too high (for then global opportunities simply do not arise) nor too low (in which case there is too little social structure for the process to take off). And if globalization does arise, we show that it often occurs abruptly and consolidates as a robust state of affairs. We also show how it is affected by improvements in the flow at which information travels in the network, or by the range at which the social network can be used to discipline behavior.

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 modern 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 response 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?* 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 it in terms of some measure of social capital that can be accumulated and maintained? Is the there a role for policy? Are there temporary policy measures that achieve permanent results?

Our model addresses these questions and provides some novel insights into the phenomenon. To end this short Introduction, we summarize 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). Furthermore, the transition is robust, not crucially depending 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 the result of a high enough accumulation of social capital, manifested in terms of dense and distant connections. This opens up the possibility of multiple equilibrium outcomes: due to history, chance, or policy some economies may heave proven successful while others not, under the same environmental circumstances. In fact, suitable policy measures impinging temporarily on the may explain such 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, decomposed in three parts. Firstly, Subsection 4.1 discusses 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 as well as some other details of our theoretical and numerical analyses 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, a tension arising between the advantages of interacting with far-away agents and the limits to this interaction imposed by geographical distance. Next, we outline these papers and contrast our different approaches.

The model proposed by Dixit (2003) can be succinctly 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, the intuitive conclusion obtains 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, assuming 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 atruistic component in payoffs whose size falls with the distance to the partner. Each individual plays the game only *once*. This allows the analysis to dispense with the information-spreading assumption of Dixit's model that is tailored to the fact that agents are involved in repeated interaction. Instead, the distinguishing characteristic of Tabellini's model is that agents' preferences (associated to the rate at which the warm-glow term 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 (i.e. the detection and reversion of cheating) 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 is quite direct: the ability to sustain cooperation (either through observability or altruism) is taken to decrease in such *exogenous* distance. In our case, instead, the relevant distance in this respect is *social* and *endogenous*, for it is determined by the evolving social network. It is precisely in the evolution of the social network that geographic distance plays a role: 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.

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 designed as truly systemic. That is, the postulated measures of globalization are based on the individual characteristics of the different "agents" rather than on their interplay with 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, multistep 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 undertaken a step in this direction. First, that paper introduces an operational counterpart of the measure of integration following from 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 again rely on the usual control variables that are considered in the growth literature but, most crucially, address the key endogeneity problem that lies at the core of the empirical issue. We find that our measure of integration is a robust and very significant explanatory variable, which 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 (when compared with the usual "local" approach) a systemic perspective that is rich and novel. We refer the interested reader to the aforementioned companion paper for details.

3 The model

The formal presentation of the model is divided in two parts. First, we describe the underlying (given) 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 the model. In this setup, agents are assumed to be involved in a collection of repeated games of uncertain duration in which cooperation is supported (at equilibrium of an 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 underlies the creation of new links, while the second one leads to 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 complementary skills of some other agent j. From an *ex ante* viewpoint, the probability $p_i(j)$ that the agent required by i's idea is some particular j is taken to satisfy:

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

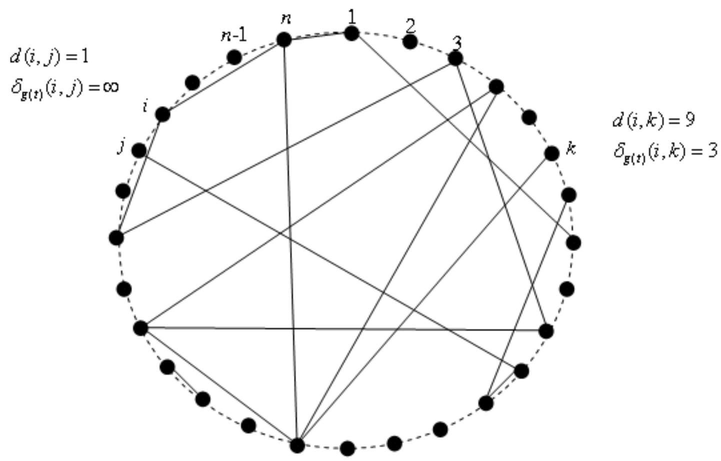


Figure 1: Snapshot of a situation at some t. By way of illustration, note that whereas the social distance is maximal (infinity) between agents i and j who are neighbors on the ring (i.e. their geodistance is the minimal of 1), the same comparison for i and k yields the opposite conclusion, namely, their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

i.e. it decays with the geographical distance (geodistance, for short) between i and j at some rate $\alpha > 0$. Since this abstract formulation admits a wide variety of concrete interpretations, in order to fix ideas we shall discuss just two of them. A first one is the most simplistic. It conceives i and j as actually meeting physically at the time i has the idea, so decay just reflects the fact that closer agents meet more often than distant ones. A second alternative is based on the idea that fruitful collaboration not only requires some 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 notion that further-away agents share less of a common ground on which such a compatibility must be based. Whatever the interpretation, the strength of these considerations is modelled by the exponent $\alpha > 0$, which for the sake of conciseness we shall refer to as (geographical) *cohesion*.

Consider then an agent i who has an idea for a project that needs the concourse of another 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 another 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."[1]

Finally, Condition L2 embodies the restrictions to be satisfied in the formation of any new link. 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.) 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 part (b) of L2 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, and hence we choose to refer to it very generally as the (quality of the) *institutions* prevailing in the economy. Within the microfounded scenario described in Appendix A, the parameter μ will be interpreted as the so-called "circle of trust" – i.e. how far on the social network individuals respond (at equilibrium) to a violation of the cooperative norm on a third party by punishing the deviator. [2]

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 all be channelled through the mechanism of link formation alone.[3]

[1] 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 skewdness in the distribution of innovation that is observed in the real world – see, for example, Jaffe and Trajtenberg (2002).

[2] Conceptually, μ can be taken to modulate 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).

[3] 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 disappear 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 (see Appendix B for details), 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. 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 understanding of what changes occur in the economy when it turns global and what are the main mechanisms that lead to 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:

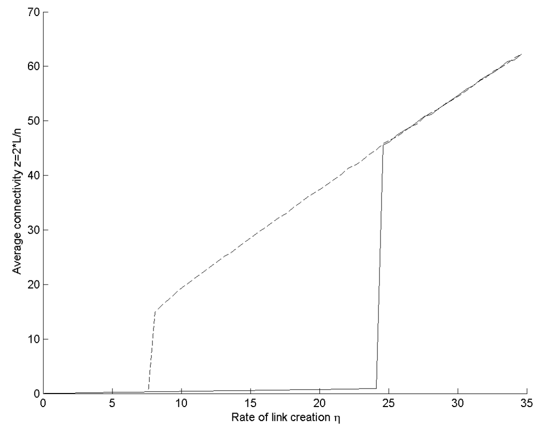
(a) average degree, as given by $\frac{2L}{n}$, where L is the total number of links and n is the population size;

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

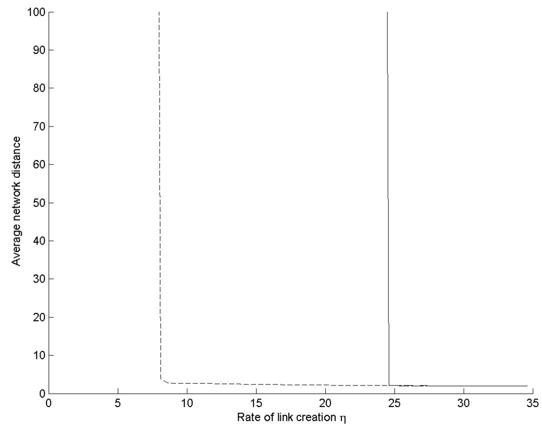
(c) the average geographical distance, obtained as $\frac{1}{L} \sum_{ij \in g} d(i, j)$;

(d) the fraction of agent pairs who are given a link-creation opportunity and actually form a new link.

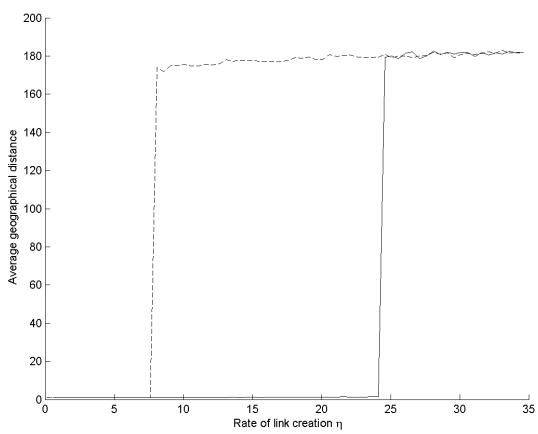
In each of the four diagrams there are two loci of steady-state values, corresponding to the sole two (stable) steady-state configurations that arise: one corresponding to initial conditions where the social network



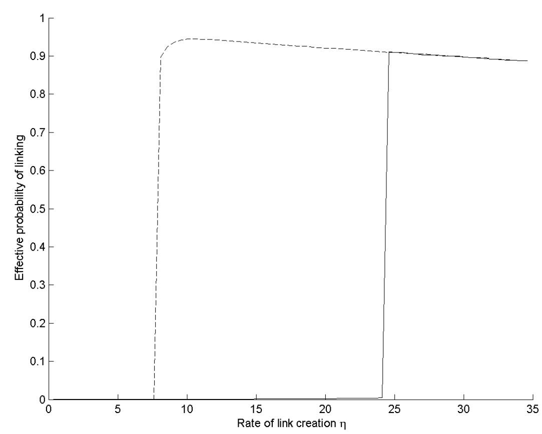
(a) Average network degree



(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 $\lambda = 1$, $n = 1000$).

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

Let us first focus on Diagram (a) of Figure 2. The two steady-state loci displayed in it 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. In terms of the parameter η there are three different regions: a low region with a unique steady-state regime of low connectivity; a middle region where the two regimes (with high- and low-connectivity) are possible, depending on initial conditions; and a high region where again there is a single steady-state regime, in this case with high connectivity. A further interesting implication of this behavior is that, if one traces the effect of gradual changes in η on the steady state (from initial conditions defined by the previous 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. [4]

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), now pertaining to average social distance (per agent pair) and average geodistance (per existing link). We find that, corresponding to each transition to 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 globalization. The latter transitions, in other words, may be seen as the other side of the coin of the former one, i.e. they must go hand in hand. Why is this the case?

To answer that question, we focus on Diagram (d), which shows that what is the main implication of globalization: to induce 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, which in turn is needed if a high-connectivity configuration is to be sustained as a steady state. (Note that if the economy enjoys many links, many are also destroyed and hence many need to be continuously created to compensate for that.)

As indicated, the behavior displayed in Figure 2 is obtained in a scenario where the level of geographical cohesion ($\alpha = 0.5$) is relatively low 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 $\alpha = 2$ (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 different diagrams and curves displayed in them.

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

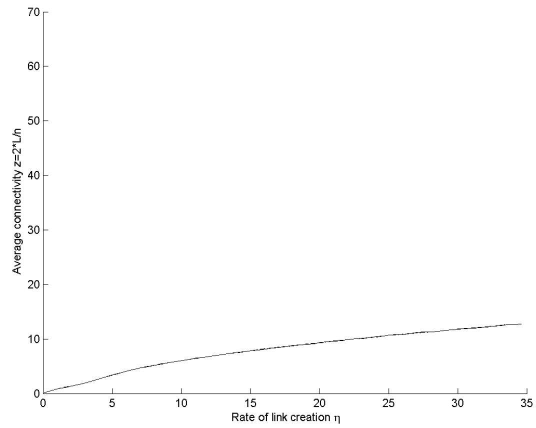
(i) Steady-state behaviour is unique throughout.

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

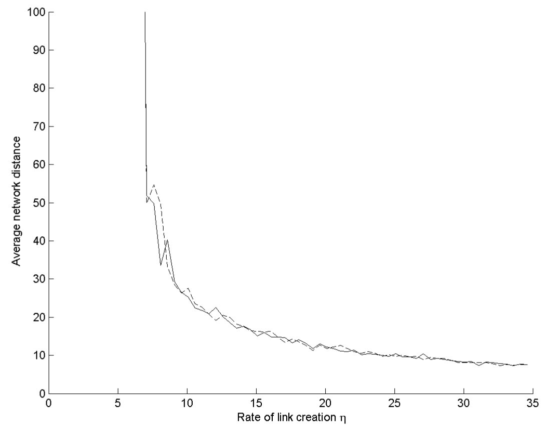
(iii) The effective linking probability remains well below unity, decreasing with η for most of its range.

Thus the first interesting observation, stated in (i) above, is that initial conditions do not matter any more 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 the same as the one achieved from initial conditions displaying a high connectivity. The key reason for this has to do with the useful structure supporting the linking process when the matching process displays significant geographical cohesion. In this case, 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. On average, therefore, it will be quite likely that there is a short network path between potential partners, who can then form a link if they do not yet have one (recall L1-L2 in Subsection 3.2.1). Under

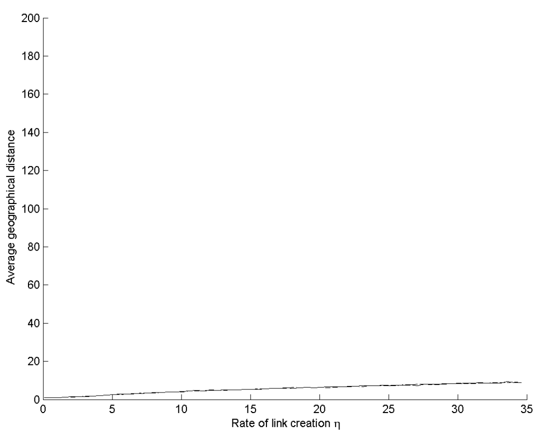
[4] This, of course, is simply a reflection of the dependence of initial conditions that occurs in the middle η region. The implications of it for policy are potentially important, as briefly outlined in Section 5.



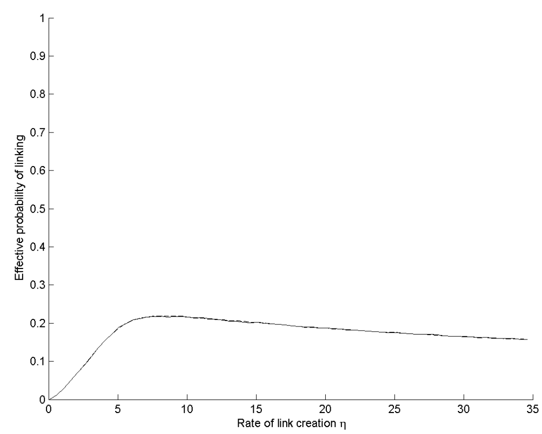
(a) Average network degree



(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 $\mu = 3$ ($\lambda = 1$, $n = 1000$).

these circumstances, initial conditions are no longer relevant to predict where the process will end up in the long run – only the value of the underlying parameters matters.

Essentially the same idea underlies point (ii) above. That is, even if η is low and the induced network is sparse, the geographic structure provided by a relatively large α allows increases in the innovation rate to translate into corresponding increases in connectivity of a comparable magnitude. But now comes an important trade-off: such higher cohesion comes at a cost, which is the point highlighted by item (iii) above. Specifically, congestion becomes an important issue, resulting from the saturation of the intensive margin in link creation – i.e. the impossibility of having more than one link per agent pair. This, 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 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 in order to build a dense social network and the detrimental congestion effects that are brought about by too much such cohesion. Indeed, our analysis will show that such a tradeoff is a general feature of the model, and that the induced tension is 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 forces explained at an informal intuitive level in the preceding subsection. We shall see, for example, that there are solid theoretical reasons why a certain threshold on geographical cohesion ($\alpha = 1$ in the benchmark limit scenario) marks the point above which the economy is able to build a dense social network from scratch. Another theoretical prediction will be that, in those cases where cohesion is not too strong, the transition towards globalization must be very sharp and locally irreversible, i.e. displays hysteresis.

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 $\phi{\cdot}\eta{\cdot}n$ 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.[5] 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 in the lage-population limit 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 $z(t)$.[6] This leads us to postulate a function $\Phi : \mathbb{R}_+ \to [0, 1]$ specifying

[5] Also remember that the number of links is *half* the total degree because every link contributes to the degree of its two nodes.

[6] 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 each of them displays polar features in a number of key respects (e.g. in terms of the variance displayed by the induced

the conditional linking probability $\Phi(z)$ associated to each average degree z, from which the steady-state condition (2) may be rewritten as follows:

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

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

As advanced, the properties of the resulting steady states crucially depend on the value of α. We 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 sense) has relatively little bearing on how new ideas arise while 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 magnitude:

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

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

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

which is the real-valued *Riemman Zeta Function*. It is well-known [7] that

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

Thus, if, $\alpha > 1$, the aforementioned probability is well-defined and 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.) Later on, we shall build upon this result to formulate the properties on the function $\Phi(z)$ used to model the infinite-population context.

degree distributions), in both cases the family of networks under consideration can be parametrized by the average degree z (cf. Vega-Redondo (2007)).

[7] 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.*[8] *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 modelled as a large random network, which belongs to a family of such networks that can be suitably parametrized by their average degree z. Under these conditions, it is well known that many interesting properties of the induced networks depend on z in a "threshold" manner.[9] 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 for any given (finite) population size n. The indicated threshold property can be formulated as follows:

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

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

Based on this property, we now formulate some useful assumptions on the function Φ in (3) that specifies the conditional linking probability of a randomly selected node. Extending the previous notation, let us write this function as $\Phi(z;\, \alpha)$ in order to account explicitly for the dependence on α spelled out in Proposition 1. First, we first two contrasting assumptions on how the function behaves locally around $z = 0$ under the two polar scenarios considered before, i.e. LGC ($\alpha \leq 1$) and HGC ($\alpha > 1$).

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, $\Phi(0;\, \alpha) = \left[\zeta(\alpha)\right]^{-1} > 0$

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) and (7) 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:

[8] 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

[9] See, for example, the extense survey of Newman (2003) or the very accessible monograph by Durrett (2007, Ch. 3) for a good account of the different results from the theory of random networks that will be used here.

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

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 just a standard regularity condition that facilitates the analysis. On the other hand, A4 first 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 then simply states that, once such sufficient level of connectivity has been attained, further increases in it still preserve *some positive* linking probability. Intuitively, this captures the idea that This, in essence, implies that linking saturation is never so complete that (in a context where the population is arbitrarily large but the average node degree is finite) there must be some, say, second neighbors who are *not* current partners themselves.

Conceiving assumptions A1-A4 as a suitable representation of the system for very large populations, we now contemplate the following dynamics for its aggregate state.

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

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 interesting conclusions that follows from a simple analysis of (10). As before, all proofs can be found in Appendix B.

We start with two results on the conditions (in particular, concerning geographical cohesion) under which 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.*[10]

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.

[10] 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.

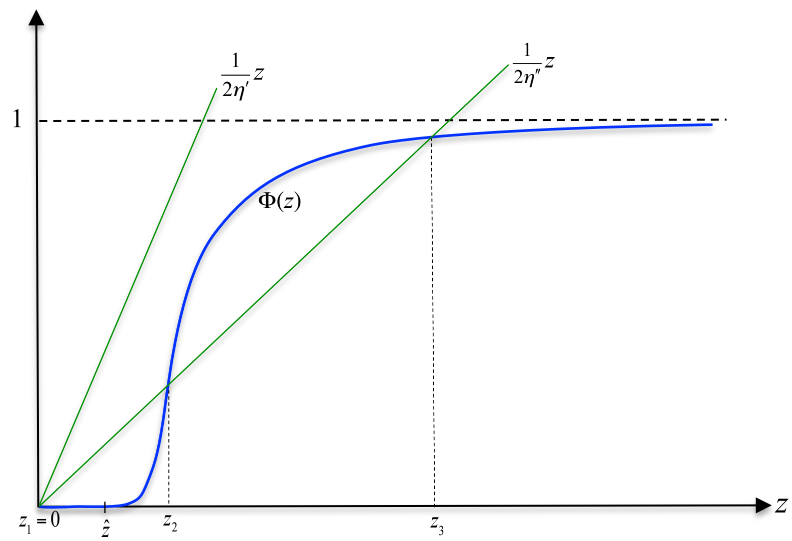


Figure 4: An illustration of how the equilibrium degree changes with η under low cohesion ($\alpha < 1$)). See the text for an explanation.

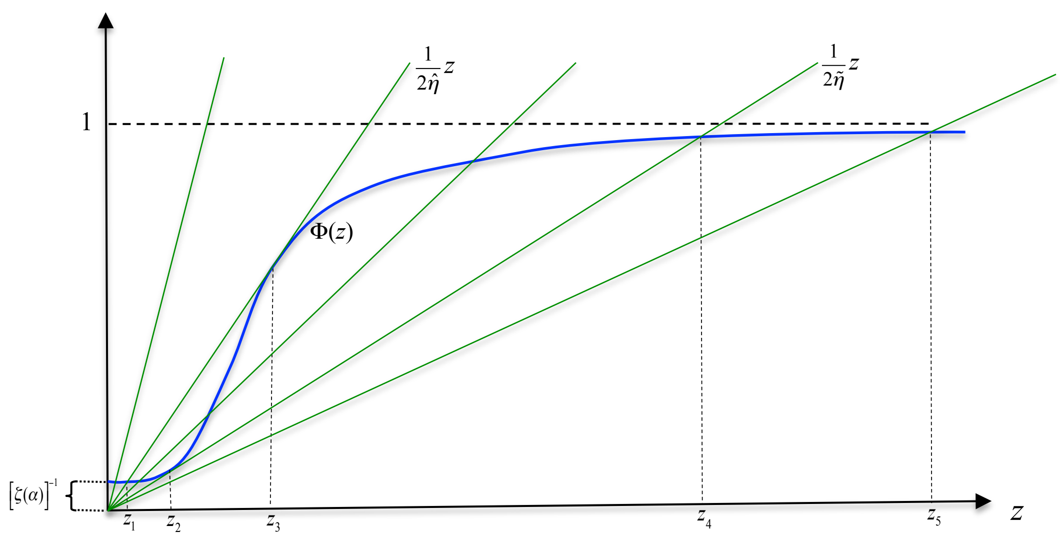


Figure 5: An illustration of how the equilibrium degress change with η under moderately high cohesion ($1 < \alpha < \bar{\alpha}$ for some relatively low $\bar{\alpha}$). See the text for an explanation.

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

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 $z = 0$). If η is high enough, such an equilibrium degree is low and remains so as long as η is below a certain threshold ($\hat{\eta}$ in the figure). But as η 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 further, the degree displayed by the corresponding equilibrium grows unboundedly (i.e. up to z_5 and beyond). As we show below, this pattern illustrated in terms of Figure 5 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. The unboundedness of the degree as a function of η, however, 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 $z^*(\eta) \geq \bar{z}$.*

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 in our analysis, the parameter μ that we have identified with the quality of institutions has played no explicit role. It is, however, implicit in the function $\Phi(\cdot)$ that models how the conditional linking probability depends on the connectivity of the network. Clearly, in a context such as that of the present benchmark model (where the population is taken to grow arbitrarily large), a significant linking probability can only be attained if μ grows with n, the population size. Under our maintained assumption that the underlying network is well-described as a random network, this is a moderately demanding requirement. For, as it is well known,[11] the diameter of the largest (giant) component of a random network grows slowly with n, even when this component represents a significant fraction of the population (possibly close to one). More precisely, the diameter grows as slowly as (i.e. at no higher order than) $\log n$. This may then be as well the assumption underlying the function Φ (and, in particular, assumption A4).

4.3 From the benchmark set-up to the general model

The benchmark model considered in Subsection 4.2 is to be conceived as a stylized and manageable representation of a context where the population is very large. This approach has proven useful in highlighting some important features of the model, but this has been achieved at the price of underplaying other issues that are best studied in a finite-population scenario. Particularly important examples of such issues are, say, the extent to which geographical cohesion may induce local saturation of opportunities, or the impact

[11]See the references mentioned in Footnote 9

of institutions on long-run performance (recall Remark 1). Another limitation of the benchmark set-up is that, given the very general 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 builds crucially upon the conceptual and methodological insights resulting from our analysis of the benchmark model in Subsection 4.2. In particular, we shall continue to rely on the steady-state condition (2) – and its reformulation as in (3) – to determine the values for average connectivity around which the system gravitate over time. Just as before, this will allow us to single out the stable steady states and the transitions between them induced by small changes in the parameters of the model. But, to do so, irst we must develop a way of determining the function $\Phi(z)$ that gives the effective linking probabilities associated to different values of z. Since we are now looking for a concrete 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 a good estimate of the function Φ and, on the basis of it, subsequently conduct a theoretical analysis akin to that carried out for the benchmark setup. Next, we describe in detail the numerical procedure used to determine the function $\Phi(\cdot)$

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 theoretical reference. For, in effect, it specifies the "notional" rate of project 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 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. For each parameter configuration, the functions

$\Phi(\cdot)$ are determined through Algorithm P.

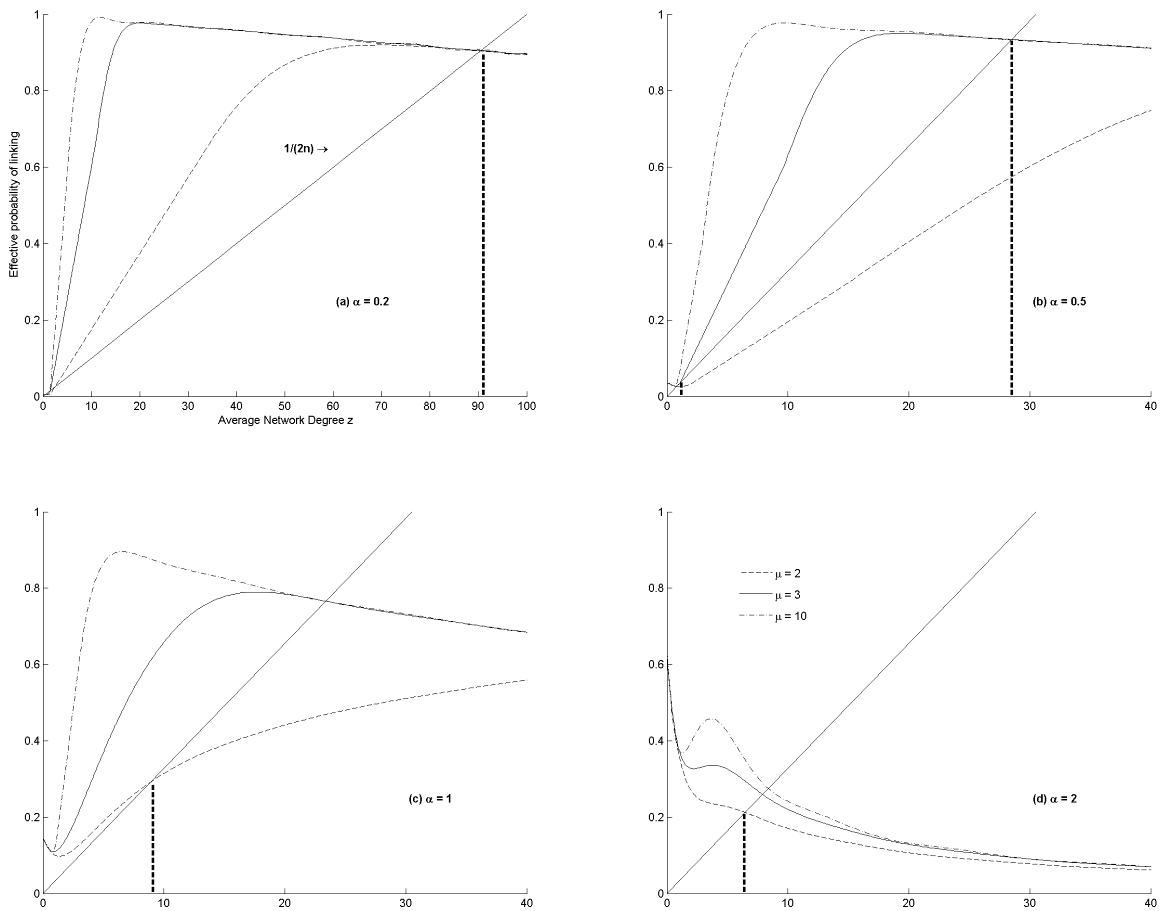


Figure 6: Graphical representation of the equilibrium condition $\Phi(z) = \frac{1}{2\eta} z$ in the general framework. The diagrams trace the steady states as points of intersection between a fixed ray of slope $1/(2\,\eta)$ and the function $\Phi(\cdot)$ computed for different institutions μ (within 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 the benchmark model. 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 $\alpha = 2$ where the differences are starkest – in which changes in η (affecting the slope of the ray) would trace a continuous change in the equilibrium values.

It is 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, attention is restricted to the results that pertain the long-run/steady-state average degree. We find that, forboth 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 confidently regard our present numerical approach as a suitable way of analyzing theoretically the model, very much along the lines used for the benchmark setup in Subsection 4.2. In the remaining part of this section, our analysis will focus on three issues. Firstly,

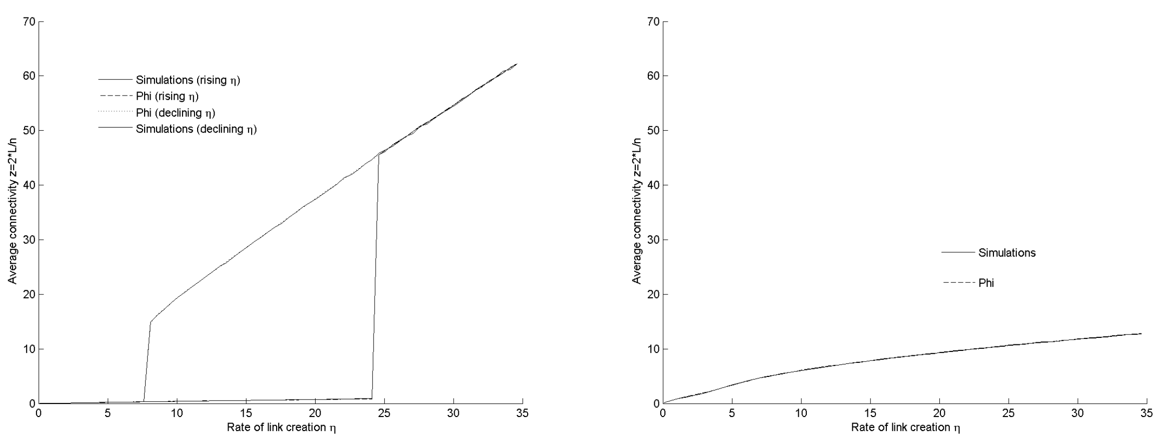


Figure 7: Comparison between the theoretical predictions and the numerical simulations on the steady-state average degree for each of the two scenarios ($\alpha = 0.5$, 2) considered in Subsection 4.1.

we shall revisit the relationship between cohesion and the rise of globalization. Secondly, we shall 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. [12]

[12]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. 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 considered 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, can only be strictly true in the limit case of an infinite population, as clearly illustrated by the simulations reported in Subsection 4.1. In general, for a large but finite population, it is to be expected 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 the lower is geographical cohesion.

Our present general approach permits the exploration of the former conjectures in a systematic manner, going beyond the stylized insights derived from our benchmark model. The results of this endeavor are presented in Figure 8 for a representative range of the parameters α and μ. In its different diagrams, 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 cohesion that 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 in line with our intuitive understanding of the model and provides a clear-cut confirmation of the suggested conjectures.

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. (Recall Footnotes 6 and 9.) This implies that, at the point where globalization occurs, μ can be no lower than the network diameter. Consequently, further improvements in institutions should entail no further formation of links. On the other hand, the reason why such a threshold should be low relative to population size is less apparent. In part, this is reflection of another prominent property of random

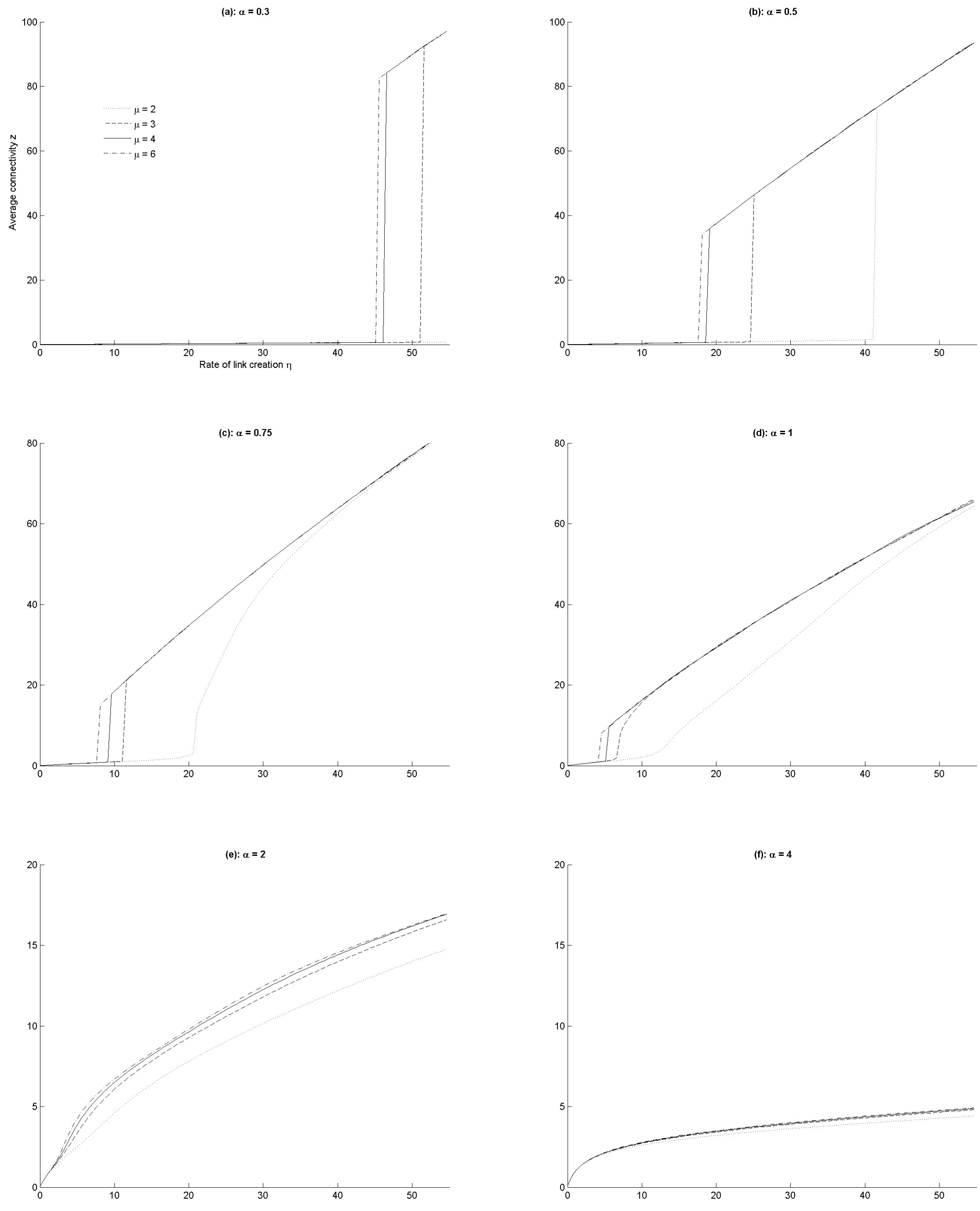


Figure 8: Numerical solution for the lowest average network degree that can be supported at a steady state, as the innovation rate η rises, for different given values of geographical cohesion α and institutions μ (with $\lambda = 1$, $n = 1000$).

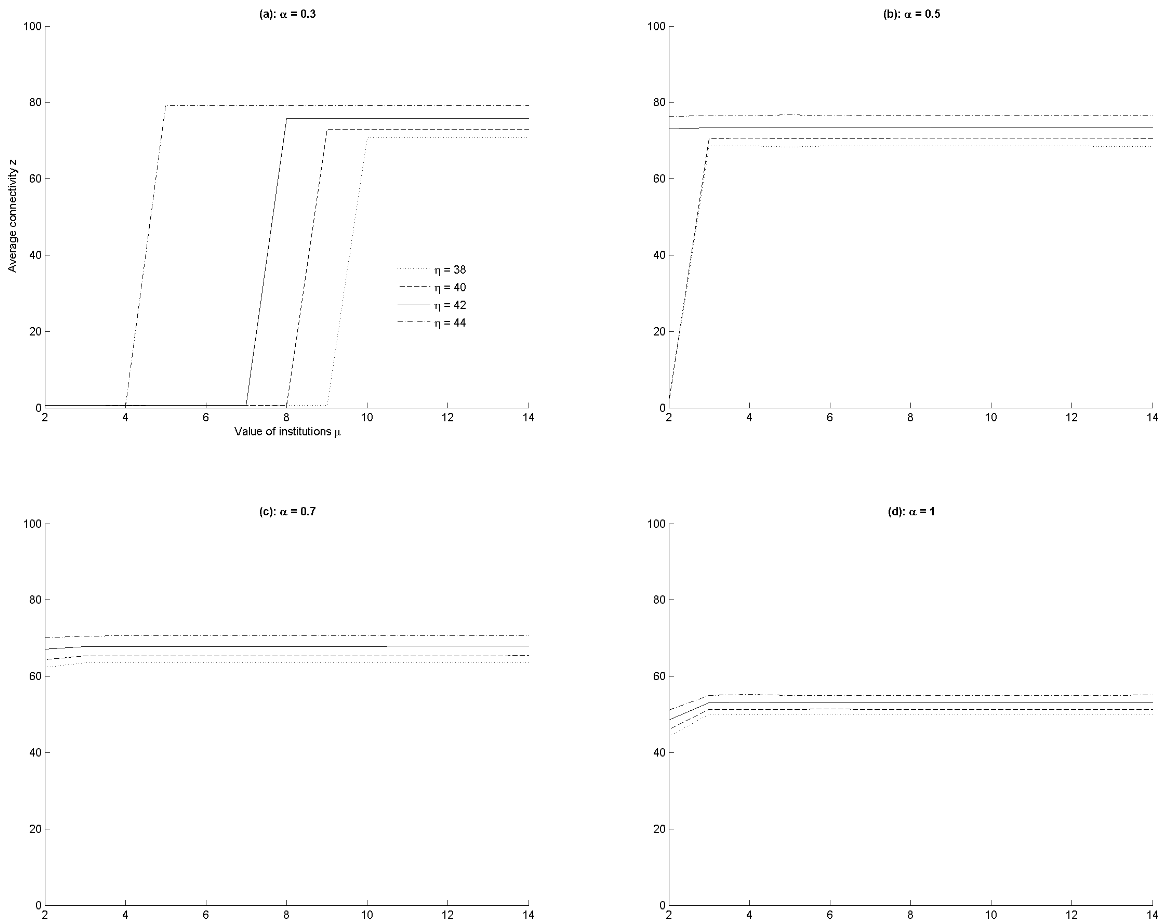


Figure 9: Numerical solution for the lowest average network degree that can be supported at a steady state, as institutions improve, for different given values of the innovation rate η and geographical cohesion α (with $\lambda = 1$, $n = 1000$).

networks that was already discussed in Remark 1: the diameter of random networks grows very slowly (i.e. scales logarithmically) with population size. Thus, the minimum value of institutions required to *sustain* a globalized state is indeed small relative to population size. What Figure 9 shows is that, in addition, the value of institutions above which such a globalized state *arises* also grows very slowly with population size.

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 do not allow such a network to arise, the rate of link creation is vanishing small. Instead, at the threshold where an institutional improvement does lead to globalization,

μ cannot lie below the network diameter. Such a threshold, therefore, also represents the saturation point beyond which any further increase is ineffectual.

4.3.3 The optimal level of geographical cohesion

Throughout our discussion, we have stressed the two contradictory implications of geographical cohesion: it provides useful structure to the linking process but, on the other hand, exacerbates the problem of congestion/saturation of *local* linking opportunities. This is why we have we expect that, in general, there will be an optimal resolution of this trade-off at some intermediate level – what we label as the Optimal Geographic Cohesion (OGC), identified with the optimal value $\alpha = \alpha^*$ that maximizes the long-run average network degree when the process starts from a sparse network. We have informally conjectured 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 evidence described 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, the lower geographical cohesion the economy the better, for it will consequently minimize the detrimental effects of congestion. Heuristically, therefore, one may think of the OGC as the lowest value of α that allows the build-up of a dense social network. Since such a build-up becomes easier to perform the better the underlying environmental conditions, the negative dependence of OGC on η and μ readily ensues.

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 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

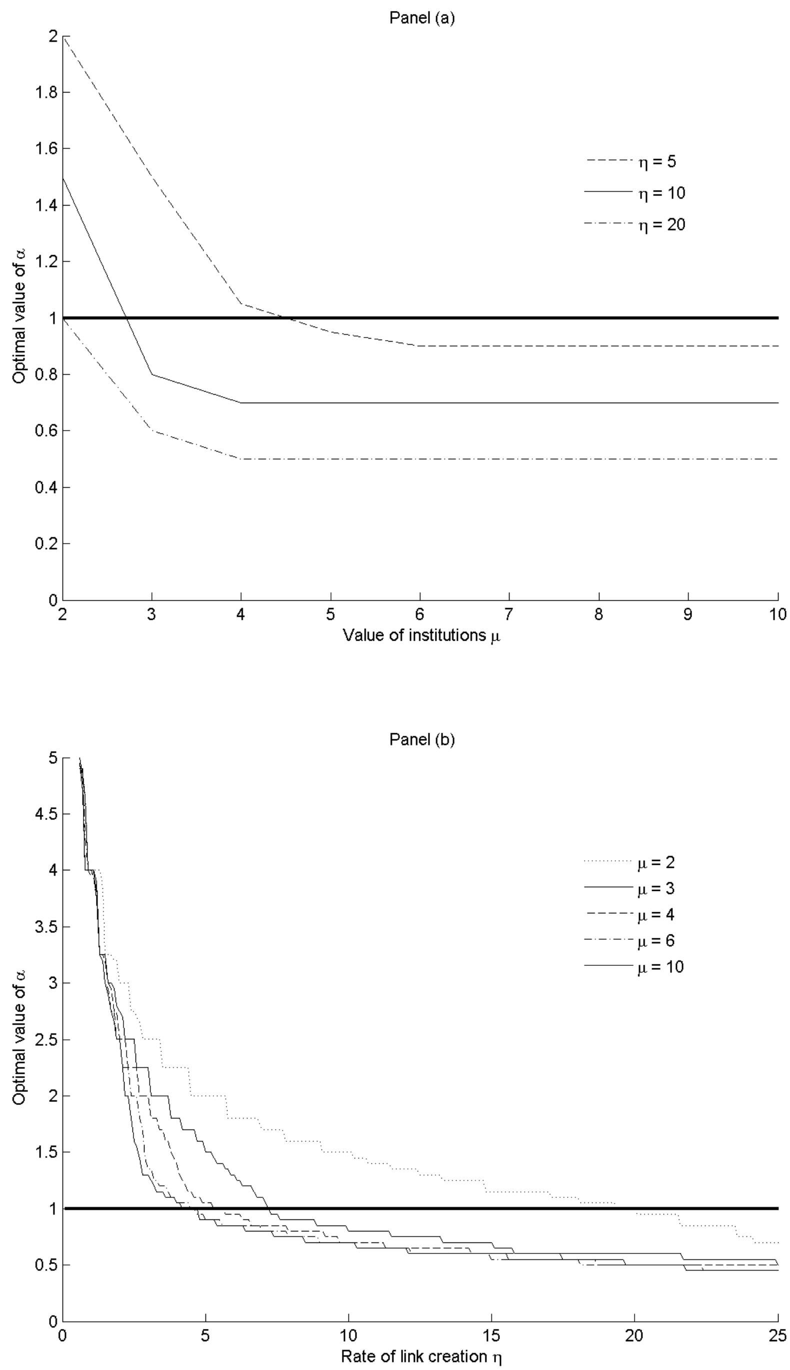


Figure 10: Optimal level of geographical cohesion (OGC), as a function of institutions μ, given different values of η (Panel (a)) and as a function of the the innovation rate η, given different values of η (Panel (b)).

(2003) – cf. Section 2.

The theoretical framework also opens the room for the discussion of policy. Just as a brief illustration at this point, recall that (if geographical cohesion is not very strong) the model yields significant equilibrium multiplicity – both in terms of the parameter region where it occurs as well as in the gap in economic performance entailed. Such multiplicity, in turn, brings in some hysteresis in how the model could react to small and temporary parameter changes. Specifically, slight changes in some of the parameters (e.g. a small stimulus to innovation) may trigger a major shift to globalization that, once attained, would remain in place even if the original environmental/parametric conditions were restored. In essence, what such policy would do is ride on the self-feeding effects that are inherent to network phenomena to attain major and robust (i.e. locally irreversible) changes in the state of the system.

Finally, another important route to pursue is of an empirical nature. As discussed as well 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 and stresses the importance of the indirect connections afforded by prevailing economic projects to support the creation of further projects. As briefly outlined in that earlier section, Duernecker, Meyer, and Vega-Redondo (2010) have built upon these insights and the corresponding theory to construct network measures of economic interacion and relate them to different measures of economic performance.

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) that 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 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 operation phase as the repeated play (with a stopping rate of $\lambda = 1$) of a simultaneous stage game displaying the following payoff table:

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

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:[13] In the former case, they incur a relatively low cost K_0 (say, because they take advantage of their geographic proximity), while otherwise the fixed 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 game.

Population-Embedded (PE) Scenario: This corresponds to a situation where the behavior in the bilateral interaction of two players can be dependent on (and also affect) the behavior displayed in some other 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 neigbors 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, if there were a continuation equilibrium in the second phase of their relationship where the induced payoff consequences can be made severe enough. This amounts to postulating the following condition:

[13] 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.

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

Under the above condition, link creation decisions could be supported bilaterally, hence rendering the problem independent of network considerations. So let us start by assuming that this condition is satisfied for 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),[14] 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 any network g, denote by $\delta_g^i(j,k)$ the network distance between some j and k when the paths that go through i are excluded. Then, given the parameter $\mu \in \mathbb{N}$ and any link $ij \in g$, define $\mathcal{N}_{ij}^{\mu}(g) \equiv \{k \neq j :$

[14] 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 pertubations from those values can be implemented that lead to the desired inequalities.

$ik \in g \wedge \delta_g^i(j,k) \leq \mu - 1\}$ as the set of those neighbors of i different from j that are no farther away from j than $\mu - 1$, along paths that exclude i. Of course, in general, this set could be empty. But if it is not, a particular agent in this set, $\varsigma_{ij}^{\mu}(g(t))$, 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 a link ij.

Relying on the previous notation, we posit that the population displays a symmetric profile where every agent i 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. 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, $\varsigma_{jk}^{\mu}(g(t)) = i$).
- 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 neighbors or $\delta_{g(t)}(i,j) \leq \mu - 1$.

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) and, on the other hand, are 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, in the network-formation process, agents will be ready to form new links with neighbors and all other agents who are not more than μ links apart, as indeed 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.[15] The essential idea here is that μ measures how strongly agents "internalize" the social norm of cooperation. If the support of this norm – i.e. the quality of institutions – is weak, agents should only be prepared to punish others if they have defected

[15] 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 aconcrete set of people with whom one shares a sense of belonging).

upon individuals who are socially very close. Instead, if institutions are strong, any violation of the social norm that agents are acquainted with is readily punished.

Appendix B
Proofs

Proof of Proposition 1

First, in order to establish 7, note that a necessary requirement for any agent ı 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 belong 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 inequality follows since the maximum probability with which any agent j can be chose 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 neighors 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 $\mathcal{R}_\beta \equiv \left\{(z, y) \in \mathbb{R}^2 : y = \beta\, z,\ z \geq 0\right\}$. 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.[16]

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-exisiting 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.

[16] 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 $m = 2, 3, \ldots$ 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 $1, 2, \ldots, m-1$ 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 $t = 1, 2, \ldots \bar{t}$, 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] 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.

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

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

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

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

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

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

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

[15] 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.

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

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

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

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

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

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

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

[23] 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.

[24] 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.

[25] 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]]

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

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

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

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

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

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

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

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