FACULTY OF LAW, ECONOMICS AND FINANCE

CREA Discussion Paper

2015-19

Management

Center for Research in Economics and Management
University of Luxembourg

Desperately Seeking Small Worlds in Corporate Boards: International Evidence from Listed Firms

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Malika Hamadi, CRENoS, University of Sassari
Andreas Heinen, THEMA, Université de Cergy-Pontoise
Nicolas Jonard, CREA, Université du Luxembourg
Alfonso Valdesogo, Department of Economics, Universidade Federal Fluminense

November, 2015

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie

UNIVERSITÉ DU LUXEMBOURG

Desperately Seeking Small Worlds in Corporate Boards: International Evidence from Listed Firms

Malika Hamadi[*]
CRENoS, University of Sassari
Andréas Heinen[§]
THEMA, Université de Cergy-Pontoise
Nicolas Jonard[¶]
University of Luxembourg
Alfonso Valdesogo[‖]
Department of Economics, Universidade Federal Fluminense [**]

November 17, 2015

Abstract

This paper analyzes the structure of national corporate board networks of listed firms in a large cross-section of countries. We introduce an explicitly bivariate nonparametric hypothesis test for small worlds, based on the comparison of observed distance and clustering with simulated measures obtained from a number of increasingly stringent bipartite counterfactuals. Using our test, we find little support for the small world hypothesis regardless of the counterfactual. Moreover, we show that results are sensitive to the choice of counterfactual. We further identify the role played by bicliques, small densely connected subsets of the network, in the rejection of the small world hypothesis.

Keywords: Boards of directors, small worlds, bipartite graphs, testing, bicliques.
JEL Classification: G34, D85, C63.

[*]School of Management, University of Leicester, LE1 7RH Leicester, UK. Email: mh463@leicester.ac.uk.

[§]Address: 33 Boulevard du Port, 95011 Cergy-Pontoise, France. E-mail: Andreas.Heinen@u-cergy.fr.

[¶]162 A, avenue de la Faïencerie, L-1511 Luxembourg. E-mail: nicolas.jonard@uni.lu

[‖]Campus do Gragoatá. Rua Prof. Marcos Valdemar de Freitas Reis, s/n, Bloco F, Gragoatá, Niterói, RJ-24210-201 (Brazil). E-mail: avaldesogo@id.uff.br.

[**]We acknowledge financial support of the National Research Fund, Luxembourg through CORE research project C08/IS/13. Experiments presented in this paper were carried out using the HPC facility of the University of Luxembourg. The usual disclaimers apply.

1 Introduction

Small worlds, defined as sparse networks consisting of densely inter-connected groups of nodes linked to one another via a few bridging ties seem almost ubiquitous nowadays in social systems[1], but also in a variety of natural and technological systems[2] (see the recent summaries of small world findings in social science and management provided by Uzzi, Amaral & Reed-Tsochas 2007, Phelps, Heidl & Wadhwa 2012). While the idea of a small world started with the Milgram (1967) experiment and gained popular acceptance with the notion of "six degrees of separation", it is Watts & Strogatz (1998) who explicitly formalized the concept. They characterize a small world as a network in which average distance is similar to that in a comparable random network but clustering, defined as the proportion of the neighbors of a node who are neighbors of each other, is present in significant excess. The value of small worlds, it is widely accepted, resides in their ability to cumulate the benefits of clustering and short path lengths. On one hand, clustering provides local overlap, which is good in general for exploitative activities requiring triangulation, repetition, validation, and maturation. On the other hand, short path lengths are useful for exploratory activities requiring novelty, exposure to differences, recombination, and rapid change. As Gulati, Sytch & Tatarynowicz (2012) put it in a recent analysis of the computer industry alliances network, "...not only are small-worlds now considered ubiquitous social structures but they are also increasingly recognized as robust drivers of individual and collective action." (p. 449).

The problem with the formalization by Watts & Strogatz (1998) is that it makes the notion of a small world contingent upon the provision of a "comparable" random network, which retains some features of the network under scrutiny, while leaving

[1] Besides boards of directors, well-studied examples of bipartite networks include movies and actors (see e.g. Watts & Strogatz 1998), co-ownership networks of companies (see Kogut & Walker 2001) as well as collaborations of scientists on papers (see e.g. Grossman & Ion 1995, De Castro & Grossman 1999, Batagelj & Mrvar 2000, Newman 2001*a*, Newman 2001*b*, Newman 2001*c*, Goyal, van der Leij & Moraga-González 2006).

[2] Watts (1999) and Newman (2001*a*) find small worlds in the world-wide web, the power-transmission grid of the western United States or the neural network of the nematode C. elegans.

others random. In that regard, the identification of small worlds suffers from the "joint hypothesis problem"[3], and the small world hypothesis cannot be tested in isolation, since it is necessarily a joint hypothesis of small worlds with a given counterfactual. Unfortunately there is no unique way of choosing such a counterfactual, or null model, and, as we show in this paper, different counterfactuals can lead to different conclusions, as to whether a particular network is a small world or not. Moreover, even with a valid counterfactual, the identification of small worlds should be based on a statistical test. Instead, the extant literature often relies on casual statements about the extent to which the network under scrutiny and the counterfactuals look 'alike' or display values that are 'not-too-different' for some structural properties. Without a proper counterfactual and a systematic comparison procedure, it is impossible to distinguish properties that are surprising in a given network from those that are to be expected in any such network. Research on antecedents and consequences of small worlds can only be conducted, once this initial identification problem is solved.

In the present paper, we analyze the national corporate boards of directors networks of all listed firms in a large cross-section of countries. We employ a rigorous testing methodology, which relies on simulating appropriately defined counterfactual networks. To this end, we introduce a formal explicitly joint nonparametric test of the small world hypothesis, based on distance and clustering, which leads us to reject the hypothesis for most countries. Our results highlight the importance of controlling for the variability and correlation of the theoretical measures, and illustrates the fact that a simple comparison of empirical measures with their average theoretical counterparts is not appropriate. We further show that a number of small but highly connected subsets of the bipartite network play an important role in the rejection of the small world hypothesis. Our paper makes four main contributions.

First, we analyze the whole universe of listed firms in a cross-section of 53 coun-

[3]See e.g. Fama (1970), Fama (1991), Malkiel (2003), who discuss how tests of market efficiency require taking a stand on an asset-pricing model

tries, characterized by different institutional frameworks, financial systems, stock markets, and traditions of corporate governance. In contrast, previous studies mostly focus on a smaller sample of firms in one or two countries: Davis, Yoo & Baker (2003) analyze Fortune 500 companies in the U.S., Robins & Alexander (2004) look at the top 250 companies in the U.S. and Australia and Conyon & Muldoon (2006) focus on the U.S., the U.K. and Germany. We believe that the unprecedented breadth of this data set contributes to making our findings both reliable and general.

Second, the paper contributes to the methodological discussion on the identification of network structures. Like earlier research, we ask whether a random process of board affiliation could explain the patterns observed in the data. We do this in a systematic way and we examine both the properties of the original bipartite network and of its projections. Our focus is on clustering and average pairwise distance, which are at the heart of the notion of small worlds. In order to address the problem of the choice of counterfactual, we conduct a large simulation experiment in which we generate realizations from three different bipartite counterfactuals with increasingly stringent restrictions. Each counterfactual corresponds to a set of characteristics that we impose on the generation of random graphs. The first one holds the size and the density of the data fixed, while the second also matches the degree distribution of the boards, and the third matches both the degree distributions of boards and directors. For all three counterfactuals we compare network measures computed from unipartite projections of both the data and our sample of simulated counterfactual networks. We therefore obtain a distribution for every network measure, against which we can compare empirical measures. This permits a principled discussion of whether random affiliation is able to replicate observed structures.

How we do the comparison is the third contribution of the paper. We develop non-parametric tests for the two hypotheses that we are interested in: random affiliation and small worlds. By construction these tests work regardless of the distribution followed by the measures computed from simulated counterfactual networks. In particular they do not rely on the assumption that these measures are normally distributed. As the small world hypothesis is in fact a non-standard joint hypothesis, which combines an equality with an inequality under the null, we rely on Kodde & Palm (1986) to design a nonparametric test for that specific purpose. To our knowledge, no previous study has strictly formulated the small world hypothesis as the joint constraint for a network to have significant excess clustering and non-significant differences in average path length in comparison with the suitable counterfactuals. Our tests can thus be seen as providing an overall metric of adequacy between data and counterfactual, which takes into account the variance of average distance and clustering and the possible correlations between them under the null. Performing a joint test is essential in order to properly control size, i.e. the probability of rejecting a true null hypothesis, since the outcome of a test with incorrect size is misleading. Our test improves on the methods currently used in the literature for the purpose of identifying small worlds, such as for example the small world ratio of Kogut & Walker (2001) and Uzzi & Spiro (2005). Our results are strongly at odds with the findings reported in the literature to date: when put to a strict statistical test only few national board membership networks qualify as small worlds. However, this does not mean that the networks are formed randomly, since we also strongly reject the hypothesis of random affiliation in most countries.

Finally we argue that bicliques in the bipartite network are responsible for the difference between the counterfactual networks and the data and for the massive rejection of the small world hypothesis. Bipartite cliques (bicliques) consist of a set of boards and directors (at least two of each type) such that all directors sit on all boards. When large enough, such bicliques are a major source of redundant links (multiple memberships) and thus a strong contributor to explaining the discrepancy between data and unipartite projections, especially in terms of distance.

2 Related literature

2.1 Identification of small worlds in bipartite networks

To identify small worlds, one needs a proper counterfactual and a test that compares the statistical properties of this counterfactual with the properties of an observed network in a systematic way. While the Watts & Strogatz (1998) formalization of small worlds is fundamentally a hypothesis on unipartite networks, boards of directors form a bipartite network. A unipartite network has nodes or actors of a single type, such as firms, inventors or directors, with edges representing for instance the transmission of information or influence. On the other hand, a bipartite network has two types of nodes, and ties only exist between nodes of different types. In general it is useful to conceive of bipartite networks as capturing project membership, where a project can cover any type of association. In board and director networks, ties between directors and boards indicate board membership.

The process which leads to the formation of boards of director networks operates on nodes of two types, and thus it is on the bipartite set of nodes that proper counterfactuals should be developed. When projecting the bipartite network on the set of participants (usually assuming a tie between two participants to the same project) membership to any given project induces a clique for all participants in that project. By construction the network of participants will be highly clustered with many more cliques than any network obtained by connecting participants at random, rather than on a project basis. This issue has been largely overlooked in the early management literature on bipartite networks, and this lends credence to the suspicion that the prevalence of small worlds identified in the literature could be partly due to false positives. Among the representative studies of bipartite networks with unipartite counterfactuals, Watts (1999) compares the intrinsically bipartite network of movie actors to the asymptotic values of path length and clustering obtained from unipartite Erdős & Rényi (1959) random networks, whose degree distributions are Poisson by construction. Moreover, this method of identification of small worlds fails to take into account the variability of the counterfactual, and relies instead on the necessarily arbitrary comparison of orders of magnitude of distance and clustering between data and counterfactual.[4]

Besides the choice of the counterfactual, another source of difficulty is the fact that, in essence, the small world hypothesis is a joint hypothesis on two characteristics of a network: path length and clustering. If we stick to the formal definition by Watts & Strogatz (1998), which has the shortcoming of being a little too vague for direct implementation, small worlds are about the simultaneous presence of almost-random distance and excess-relative-to-random clustering. In the literature, a workaround for the joint constraint has often been to compute the small world quotient, i.e. the ratio of clustering, scaled by the average clustering observed for the random counterfactual, to similarly re-scaled average distance. This small world quotient was introduced in Kogut & Walker (2001) and renamed the Q-statistic by Uzzi & Spiro (2005).[5] While the ratio has the apparent merit of condensing the two important properties that characterize small worlds into a single metric, it confounds simultaneous changes in the numerator and denominator. Moreover, identification of small worlds is carried out by comparison of the empirical Q with an arbitrary threshold, above which a network is claimed to be a small world. Thus, given the lack of an objective threshold for Q, using this workaround may result in misleading conclusions.

[4]The asymptotic formula for distance in (unipartite) Poisson random graphs essentially states that average path length scales logarithmically with network size, so $\ln(N)/\ln(k)$ where N is the number of nodes (actors) and k the average number of nodes connected to a random node (the average degree), while clustering asymptotically approaches $k/(N-1)$. This is an asymptotic average or expected value, and the quality of the approximation increases with the size of the network.

[5]Many papers use a Q-statistic based on unipartite Erdős Renyi counterfactuals to identify small worlds in affiliation networks. see e.g. Baum, Shipilov & Rowley (2003) in their study of the affiliation network of bank syndicates in underwritings, Fleming, King & Juda (2007) in their study of patenting inventors in the U.S., and even more recently Rosenkopf & Padula (2008) in their study of the mobile communications industry or Gulati et al. (2012) in a study of the global computer industry. This ratio has been used in other fields, as well, under different names. For instance in biological networks, Humphries, Guerney & Prescott (2006) define a small-world coefficient that they call σ, which is exactly the same as the Q measure of Kogut & Walker (2001).

The approach of Newman, Strogatz & Watts (2001) represents an important milestone in the identification of small worlds in bipartite networks.[6] Using this approach, Newman (2004) finds short paths and features of small worlds in scientific co-authorship networks in biology, physics and mathematics, while Uzzi & Spiro (2005) use this approximation for bipartite networks with arbitrary degree distributions in their study of hits, flops and fails in Broadway musicals. The limitation of the Newman et al. (2001) approach is that, in spite of its elegance, it remains an approximation that holds in the limit for very large sparse networks, and by construction it excludes bicliques or more generally cycles of any length. Moreover, while this approach delivers a point estimate of the counterfactual measure, it does not deliver a distribution, and thus it does not lend itself to statistical testing. The problem is that a difference that appears small in absolute value might look much larger when properly scaled by its standard deviation under random affiliation. We review more recent procedures for identification of small worlds in boards of directors networks in the next subsection.

2.2 Boards of directors and small worlds

Within the broader family of affiliation networks, governing boards have been particularly well studied. This is probably due to their unique situation at the intersection of sociology, governance, strategy, economics and finance (on the latter two, see Hermalin & Weisbach 2003, Adams, Hermalin & Weisbach 2010). Management research has examined the properties of boards and their implications from different angles, including agency theory, institutional logics, social network and resource dependence theories. There is a fairly general agreement that board members should be chosen to enhance the organization's direct and indirect access to advice, counsel, legitimacy and other resources available outside the firm (see

[6] Starting with the probability generating functions of two random variables, the degree of a random board and the degree of a random director, the approach uses power and composite functions to establish asymptotic approximations for the expected value of several quantities of interest in the unipartite projections: degree, clustering, distance and size of the largest component.

Mizruchi & Stearns 1994, Hillman, Cannella & Paetzold 2000). However, at odds with this view of directors as resource providers, Davis (1996) reports that for the U.S., board interlocks match neither inter-industry exchange relations, nor banking ties, nor ownership or control, concluding that board membership mostly reflect the social processes at work among directors: " 'who knows who' can sometimes be more important than incentives and expertise: directors join boards because they can learn from the experience, because they owe favors, or because they feel an obligation to do so." (Davis (1996), p. 156).

One phenomenon which is manifest across all countries and industries is that corporate boards are characterized by large extents of director commonality. This feature leads to the emergence of a very large connected component, which contains most of the population of firms and directors. Because of interlocking, even if they do not sit on the same board, two directors can be related through a chain of acquaintances, as long as they are part of the same connected component. This favors the dissemination of knowledge generated in one board to other boards, and eventually to the larger economy. A large body of empirical literature supports this view of interlocks as channels for the circulation of information among organizations, and documents the central role of the social network of directors in the diffusion of behaviors and strategies.[7]

Several studies also find evidence of small worlds in boards of directors networks. These studies all share the problem of how to identify small worlds. Moreover the choice of counterfactuals is seldom compelling, and practically never discussed. Using the Q-statistic based on a unipartite counterfactual, Kogut & Walker (2001) find small worlds in the bipartite cross-ownership links among German firms, in

[7]For instance Palmer, Devereaux & Zhou (1993) show the impact of interlocks on changes in corporate structure Haunschild (1993) shows how inter-organizational imitation affects corporate acquisition strategies; Davis (1991) and Davis & Greve (1997) show how the adoption of anti-takeover defenses propagates through the directors network; O'Reilly, Main & Crystal (1988) and Hallock (1997) analyze the effect of interlocks on the determination of compensation levels; Bouwman (2011) and Shipilov, Greve & Rowley (2010) show how interlocks matter for the propagation of corporate governance practices (board size, fraction of outside directors, director base pay, tenure length, CEO duality, etc.).

which two firms are linked when they own stocks in some common other firm or in each other. Using the same technique, Davis et al. (2003) find small worlds in the bipartite network of Fortune 500 boards of directors networks. As we discussed earlier, these studies might be biased in favor of finding small worlds.

Conyon & Muldoon (2006) acknowledge the bipartite nature of their data and find that the networks of U.S., U.K. and German boards of directors are closer to randomness than to small worldliness when they separately compare distance and clustering of the data and of the measure computed with Newman et al. (2001). Finally, two papers on bipartite networks use counterfactual distributions rather than point estimates. Robins & Alexander (2004) provide somewhat mixed evidence of the presence of small worlds for the U.S. and Australia, while Conyon & Muldoon (2008) find that the ownership and board of director networks of U.K. firms have path length and clustering different from their random counterfactual, with path length longer than expected from the counterfactual, and higher clustering. This leads them to reject a purely random association origin for the board networks but not to accept the small world hypothesis, which is consistent with our findings in this paper.

3 Methodology

Our testing methodology resolves a number of the challenges posed by the identification of small worlds. We propose a rigorous explicitly bivariate nonparametric small world test based on the comparison of observed distance and clustering with measures obtained from simulated bipartite counterfactuals. Our test improves over existing identification procedures of small worlds in a number of ways. First, unlike the methods that rely on loose comparisons between observed values and point estimates of measures based on a counterfactual, our test takes into account the variability of the counterfactual under the null. Second, the hypothesis testing methodology provides an objective metric to quantify the distance between the data and the counterfactual. Third, as it is bivariate and nonparametric, our test properly accounts for dependence between measures, and it delivers a correct size, regardless of the distribution followed by the counterfactual measures. Finally, unlike the Q-statistic, it properly accounts for the effects of distance and clustering, and does not allow for any sort of compensation between departures from randomness in both measures. Our testing method has more general applicability, and could also be used in joint tests of a larger set of network measures, e.g. distance, clustering, degree, betweenness, and even combinations from both projections simultaneously, to check the overall adequacy of a given observed network with a sample of counterfactuals.

3.1 Networks: definitions and measures

A network consists of a set of nodes and edges between them. Boards of director networks are bipartite, or affiliation networks with two distinct types of nodes, a set of corporate boards $B = \{B_1, \ldots, B_{N_B}\}$ and a set of directors $D = \{d_1, \ldots d_{N_d}\}$. Consider also a set g of K unordered pairs of nodes from different sets. An edge $B_i d_j \in g$ is formed in the bipartite network (B, D, g), whenever director d_j sits on board B_i, which creates a link between them. Figure 1(a) illustrates the definition, with $N_B = 3$ boards, $B = \{B_1, B_2, B_3\}$ and $N_d = 9$ directors, $D = \{d_1, d_2, \ldots, d_9\}$. Boards are displayed as grey squares, whereas directors are represented with white circles.

The *neighborhood* of board B_i consists of the nodes to whom B_i is directly connected, and is denoted $\mathcal{N}_{B_i} = \{d_j : B_i d_j \in g\}$. In the same fashion we define the neighborhood of director d_j as $\mathcal{N}_{d_j} = \{B_i : B_i d_j \in g\}$. For instance in the upper part of the figure, when considering the bipartite network, $\mathcal{N}_{B_1} = \{d_1, d_2, d_3, d_4\}$ whereas $\mathcal{N}_{d_8} = \{B_3\}$. The size of the neighborhood of board i is the number of ties held by firm i and we denote it $N_{B_i} = |\mathcal{N}_{B_i}|$, where $|A|$ is the cardinality of set A. This is called the *degree* of node B_i, and corresponds to the size of board B_i. An interlocking director has degree at least two: for instance $\mathcal{N}_{d_3} = \{B_1, B_2\}$. The *degree sequence* is, for each type of node, a non-increasing sequence of vertex degrees, and the *degree distribution* is the corresponding distribution. For boards, this is the board size distribution, for directors it corresponds to the distribution of the number of seats per director.

From the bipartite network, the board network obtains by projection on the set of boards B with the convention that two boards are connected when they have at least one common director. Symmetrically, the director network obtains by focusing on the set of directors D and adopting the convention that two directors are connected when they jointly sit on at least one board. Directors who sit on more than one board are called interlockers, and they are the ones that make the connections in the unipartite projections. For instance in the board membership network, were it not for interlocks, firms would be isolated from each other and directors would only know those directors who sit on their board.

Figures 1(b) and 1(c) illustrate the projections on boards and directors. The board projection has B_1 connected to B_2 and B_2 connected to B_3, i.e. three nodes and two edges. The degree sequence of the board projection is $\{2, 1, 1\}$ with corresponding distribution $(1, 2/3; 2, 1/3)$. When looking at the bipartite network, it becomes apparent that there are three interlocks, directors d_3, d_4 and d_6. Directors d_3 and d_4 sit on boards B_1 and B_2, and director d_6 sits on boards B_2 and B_3. There are 9 nodes in this network, and 17 edges. Each board appears as a complete clique in the director network, and interlocks are the directors who are part of more than one such clique. The degree sequence of the directors projection is $\{6, 5, 5, 3, 3, 3, 3, 3, 3\}$ with corresponding distribution $(3, 2/3; 5, 2/9; 6; 1/9)$, which reflects the fact that from the $N_D = 9$ directors, d_6 is connected to 6 other directors, d_3 and d_4 are connected to 5 other directors and the remaining 6 directors are connected to 3 other directors each. Observe that there is a multiple interlock, also referred to as biclique, in the case of boards B_1 and B_2: directors d_6 and d_7 both sit on boards B_1 and B_2. More formally, a $(2, 2)$-biclique is formed whenever two directors are members of the same two boards. This creates a redundant link, which gets lost in the board projection when done according to the simple principle of "at-least-one-common" director. The same happens symetrically in the director projection.

Any of the three networks in Figure 1 with its relevant sets of nodes S and edges g lends itself to the calculation of a number of statistics. The network measures we use are standard, all of them formally defined in Appendix A. The *average degree* is the neighborhood size, averaged over all nodes. The *density* of the network obtains as twice the ratio of average degree to the number of nodes minus one. This facilitates comparison of networks of different sizes. Moving one step beyond neighborhood, one can explore the properties of neighborhoods of neighborhoods. Here we care about local overlap or embeddedness, i.e. the extent to which a node's neighbors are neighbors of each other. The simplest and most common measure is *clustering*, defined as the number of friendship links among i's friends divided by the number of possible links among them. It ranges from 0 when the focal node's neighbors have no relationship, to one when the node's neighborhood is a clique. The *clustering coefficient* in a network is the average of that statistic taken over all nodes[8]. Taking things one step further from the focal agent, it is possible to examine the global network structure (all indirect ties in the network). The *distance* between two nodes is the number of links in the shortest path between them. *Average distance* is the average taken over all pairwise distances. This measure only makes sense when all pairwise distances are finite. To cope with disconnected networks one can restrict attention to distances between reachable pairs, or restrict attention to the largest

[8]An alternative approach to clustering is to look at the overall transitivity of relationships, i.e. the fraction of transitive triples in the whole network. Transitivity is equal to three times the ratio of the number of triangles in the network, divided by the total number of open triangles centered on any node (each triangle involves 3 open triangles, hence the factor 3). Whereas transitivity is a ratio of two sums taken over all nodes, clustering is a sum, taken over all nodes, of node-specific ratios. The two measures differ in general, but both are reliable indicators of the extent to which neighbors of neighbors tend to be neighbors of each other.

connected component, also called the *giant component*.

3.2 Construction of counterfactuals

In order to test for the presence of non-random elements in a given network structure, we build a distribution of "comparable" random networks. Only then can we analyze to what extent the observed data significantly depart from the sample of random affiliation counterfactuals. We build our counterfactuals by randomly allocating directors to boards. This results in counterfactuals that are random bipartite networks with size (number of boards and directors) identical to the data. We reiterate that generating bipartite counterfactuals which are then projected on the sets of boards and directors is the correct approach.

We test our counterfactuals using two methods that differ in the input they use to simulate from. *The first method uses the size and the degree distributions computed from the whole network*, which includes all boards and directors, regardless of whether they belong to the giant component, or not. In that case, the network statistics will be computed on giant components whose size can be different for every simulation, and that can also be different from the observed giant component. *The second method follows the tradition of generating connected random networks only using as input data of the observed giant component.*

While simulating from the entire network is relatively easy, except for its computational cost, simulation from the giant component requires more care, since we need to ensure that the random networks we generate are also connected. Failing to do so would lead to possibly disconnected graphs. Thus we construct a four-stage algorithm as follows. Our algorithm starts out as what Milo, Kashtan, Itzkovitz, Newman & Alon (2004) call a "matching" algorithms and ends up with a few rounds of "switching". First we simulate degree sequences, then we match directors and boards with these simulated degree sequences; next, when simulating from the giant component, we connect the graph, and finally we swap edges to eliminate any possible departure from randomness. For a more detailed explanation of the algorithm, see Appendix B. Our approach involves considerable computational effort, both for the simulations and the computation of network statistics, particularly in large economies such as the U.S., Canada, India and Japan, where the bipartite board affiliation network contains more than 10,000 nodes. We resort to parallel computing on a cluster of 100 machines to get the computational time for the whole experiment down to a couple of days.

We consider three bipartite counterfactuals with increasingly stringent restrictions, that we use to identify small worlds. For all three counterfactuals we compare the unipartite projections of both the data and our simulated counterfactual networks.

First counterfactual: Holding size and density fixed

The first counterfactual only imposes the same number of boards N_B, directors N_d and links K, as in the actual network. Even though this is the least restrictive possible counterfactual, it still takes as given the number of listed firms in the economy, and more importantly, the size of the pool of potential directors, as well as the intensity of connections, i.e. the number of links per board (K/N_B) or per director (K/N_d). This counterfactual makes a number of implicit assumptions.

While it can be argued that keeping fixed the number N_B of existing listed firms in a country is a very natural assumption, this should nonetheless be kept in mind, when results are used for the purpose of cross-country comparisons. Of course this assumption also depends on the target population of the study. The management literature analyzes very different sets of firms, and usually focuses on rather small sets, presumably because this makes handling the network easier and less computationally expensive.[9] While this is somewhat arbitrary, it guarantees that all the

[9]For instance for the U.S., Davis (1996), Davis et al. (2003) focus on Fortune 500 firms, Kogut & Walker (2001) study German firms, Robins & Alexander (2004) study the 200 top non-financial and 50 top financial companies in the U.S. and in Australia, Conyon & Muldoon (2006) work with a sample of 1733 listed firms in the U.S., all listed firms in the U.K. and 2354 German firms, both private and listed. Like Conyon & Muldoon (2008) for the U.K. we instead choose to focus on the

firms we include in the network in any given country are subjected to the same legislation in terms of transparency, disclosure, governance etc.

Imposing the number of existing connections K, on top of the previous assumption, amounts to fixing average board size (K/N_B). While there are certainly a number of legal requirements that have an impact on board size, such as for example laws imposing unitary or dual boards, within those constraints, the size of the board of a firm is endogenous. Indeed, board size is a conscious strategic decision of the firm, which might be viewed as an optimal response to the environment the firm is operating in. Average board size in a country results from aggregating these individual firm decisions, and thus "small worlds" studies will have nothing to say about this important feature of the network.

Finally, imposing the size N_d of the pool of directors that are allowed to serve on the boards of listed firms is the assumption that carries the most weight. Along with the number of connections K, this pins down the average number of boards that each director sits on. This is not innocuous. Essentially it means that the only eligible directors are drawn from the pool of directors who already serve on the boards of listed firms. Making the size of that pool endogenous is simply impossible, yet the size of that pool can to a large extent influence the amount of "cliquishness" found in the data. Who belongs to the pool of eligible directors and how one enters that pool can also determine other features of the observed network, but this is outside of the range of issues that a "small worlds" study can address.

Second counterfactual: Degree distribution of boards

While the first counterfactual imposes several implicit assumptions, there are also a number of dimensions in which it leaves more freedom than is desirable. For instance, if in a country the minimum legal board size is 4, there is nothing that stops the counterfactual from producing networks with smaller boards, as long as the average board size is correct. More generally the first counterfactual is unlikely

full set of listed of each country.

to be able to reproduce either the board size distribution, or the directors degree distribution.

Thus we further impose the board degree distribution on our next counterfactual. In other words, we take the observed board size distribution as given when generating graphs. While keeping fixed the number N_d of directors, we do not impose anything on their degree distribution, and we assign directors randomly to the existing boards.

Third counterfactual: Degree distributions of boards and directors

Our last counterfactual imposes the degree distribution of both boards and directors. This is our strongest assumption, and it is the one followed also by Conyon & Muldoon (2008). Imposing the degree distributions means that we remain silent about the potential determinants of these distributions, most importantly the degree of the directors. This implies for instance, that our counterfactual imposes the same number of "star" directors, who sit on a very large number of boards, as in the data, and this is a feature, whose emergence we are not trying to explain. The same holds for other features of the degree distribution of the directors.

3.3 Testing procedure

Our tests rely on the comparison of observed network measures with simulations from a counterfactual. For each national network, we generate a sample of $S = 1,000$ simulated bipartite counterfactual networks and compute the empirical distributions from the simulated sample for each of the following five measures: size of the giant component, average degree, clustering coefficient, average distance and number of $(2,2)$-bicliques. We perform essentially two kinds of tests: tests on individual measures and joint tests of several measures. For individual measures, we test whether each measure is compatible with its empirical distribution calculated from the counterfactual networks. To carry out the test we compare the observed measure to the distribution under the null of random affiliation, and reject the null hypothesis when the observed value is extreme. In other words, we form a non-rejection region for the null hypothesis at 95%, which is the interval between the 2.5% and the 97.5% quantiles of the distribution, and we reject the null whenever the observed measure falls outside this region.

An alternative that is sometimes used is to compute a z-score, which normalizes the difference between observed measures and averages over the simulations under the null by the standard deviation of the simulations. This approach of computing a standard score is followed by Robins & Alexander (2004), and suggested in Conyon & Muldoon (2008). z-scores express the distance between data and simulations in terms of standard deviations. They can only be used as a statistical test under the assumption that the distribution under the null of random affiliation is Gaussian, which is not necessarily the case. The test then consists in rejecting the null hypothesis at 5% whenever the z-score exceeds 1.96 in absolute value.

Besides tests of individual measures, the two main hypotheses that we are interested in, *random affiliation* and *small worlds*, have to be formulated in terms of the joint behavior of two network measures, average distance and clustering coefficient. Thus we also need to consider joint tests of the overall compatibility of the data with the joint distribution of the counterfactual, which take into account the correlations between the measures under the null hypothesis. If distance and clustering turn out to be very correlated, then their weight should be decreased, as the additional information that each measure brings relative to the other is low. A joint test allows us to do precisely that. Moreover, a joint test allows us to correctly control for the size of the test.[10]

A test of random affiliation corresponds to the hypothesis that the actual values of distance and clustering are both compatible with the distribution of distance and clustering observed for our counterfactual network. In other words, we want to check that the observed network is not "unusual" for the type of network we consider. If

[10]It is well-known that if tests are conducted separately on each measure, even when these measures are independent, the size will not be correct, see e.g. Stock & Watson (2010), Chapter 7.

one is willing to assume normality of the simulations, this hypothesis can be tested with a traditional Wald test statistic. Instead, we rely on a nonparametric version of the test, which does not rely on any distributional assumption. We form a non-rejection region at 95%, and we reject the null hypothesis, whenever the observed measures fall outside of this region. For more details on how we compute this non-rejection region, see Appendix C.

Next we want to devise a test of small worlds. A small world is defined by the conjunction of short average distance and local clustering. The vagueness of the definition is somewhat comfortable, and it could probably be argued at superficial glance, as it is in other papers analyzing board membership networks, that our networks are small worlds. Instead we use a strict definition of small world, namely that average distance D is not significantly different from that found in a sample of random counterfactuals while the clustering coefficient C is significantly larger. We use this definition for our small world test, which corresponds to the hypothesis that the actual values of distance is compatible with the distribution of distance, while the actual value of clustering is either compatible with the counterfactual network or greater. In other words, we want to check that distance in the observed network is not "unusual" for the type of network we consider, while clustering is not "unusually" lower than in the counterfactual.

This is not a standard test, as the null hypothesis combines an equality and an inequality restriction, but Kodde & Palm (1986) show how such a nonstandard test can be carried out under the assumption of joint normality. Essentially the data has to be projected onto the set that corresponds to both the null and the alternative hypothesis. The test statistic they propose follows a mixture of χ^2 distributions, with weights that reflect the probability of being in a regime with distance in the data above (below) distance in the counterfactuals, and the two χ^2 distributions represent the joint distributions, conditional on being in one or the other regime. We develop a nonparametric version of this test, which does not rely on the assumption of joint normality. Appendix D shows how we implement this test by comparing observed measures to a non-rejection region based on the distribution of the simulations. To the best of our knowledge, it is the first time that a joint test is used to evaluate the small world hypothesis.

It should be noted that our null hypothesis does not exactly correspond to the definition of small world, as it allows clustering to be larger or equal to simulated clustering, whereas the definition implies that clustering should be strictly larger. Strictly speaking, our null hypothesis is a combination of random affiliation (equality of distance and clustering to simulated values) and small worlds (equality of distance but clustering strictly greater than simulated clustering). Thus whenever we reject our null hypothesis, we can be very confident that there is no small world.

4 Data and results

4.1 Data set

Our data is extracted from Bloomberg's Businessweek.com website during the month of June 2010. It contains the composition of the boards of directors for all listed firms around the world and for a number of private firms as well. In our sample selection we discard private firms, because it is not clear what the coverage is for each country. Instead, we focus attention on listed firms, as this provides us with an objective criterion to constitute the sample. The data covers a total of 193 countries, but we restrict ourselves to countries where the largest component is made up of at least 20 firms, which leaves us with 53 countries. Both directors and firms are identified by a unique code that can be found in the downloaded html files. These codes allow us to construct the list of directors on the boards of all companies, as well as the list of boards that each director sits on. With this sample, we construct bipartite graphs of directors and firms for each country in our sample.

Table 1 provides descriptive statistics for the national networks: the total number of firms and director seats, the average board size and number of directorships held, and the number of firms and directors in the largest connected component. The networks we examine range from a maximum of 9249 firms and 54764 directors for the U.S. to 30 firms and 213 directors for Mauritius. Limiting attention to the largest component, the networks are significantly smaller, ranging from a maximum of 4375 firms and 30223 directors, still for the U.S. to 20 firms and 178 directors in the case of Bahrain. Not surprisingly, large economies have large corporate networks. Overall, the largest component of a national network encompasses roughly half of the firms and directors of that network.

4.2 Holding size and density fixed

The results of this first counterfactual can be found in Panel A of Table 2. The first columns give the results for the giant component when we simulate from the whole network. Out of the 53 countries in our sample, board projections show excess clustering in all but 4 countries, but we fail to reject the hypothesis that simulated distance is equal to the distance in the data in only 17 countries, and thus we also fail to reject the null of small worlds in only 19 countries. This means that we only find small worlds in less than half our sample of countries. Moreover, for 3 of them, we also fail to reject the null hypothesis of random affiliation, thus we are down to only 16 countries with small worlds, which is less than a third of the countries. For directors, we have less excess clustering, significant only in 35 countries, a slightly smaller number of countries with actual distances compatible with the simulation, which reduces the number of countries with small worlds to 16, and simultaneously increases the number of countries that are compatible with random affiliation to 11.

However, these results are based on giant components whose size is very different from the ones observed in the data. For example, for board projections, the actual giant components are larger than the simulated ones on average by 36.8%, significantly so in 37 countries, while they are smaller in only 2 countries. A larger giant component will mechanically lead to larger distances in simulation, which is what we observe, since 32 (33) countries have larger distances in simulation than in the data in board (director) projections. This follows from the fact that in random graphs, distance is of the order of $log(N)$, where N is the number of nodes in the graph.

Next, we analyze the results obtained when we simulate only with data from the actual giant component, and we impose connectivity, thus by construction there cannot be any discrepancy in size between the data and the simulations. When we simulate from the giant component, we impose that all boards and directors have a least one seat, and the rest of the links are distributed randomly between them. Qualitatively, the results are very similar to the previous case. There is still excess clustering in most of the countries, the number of countries with distance compatible with the simulations is 17 for board and 15 for director projections, and overall this produces almost the same number of countries with small worlds as before. The main differences are that the number of countries consistent with random affiliation in the directors projection decreases from 11 to 2, and the distances in the simulations are reduced, so that for most countries (32 for board and 31 for director projections) simulated distances are significantly below observed ones, whereas before the number of countries above was similar.

Robins & Alexander (2004) use a similar approach on the U.S. and Australian corporate networks and find that the small world hypothesis is supported when the median pairwise distance is used rather than the average pairwise distance over all reachable pairs (and the sample median is used rather than the sample average). In their study the actual and sample median coincide and this, together with the presence of significant excess clustering, leads the authors to accept the small world hypothesis for the U.S. and Australia.[11] Another important difference in their study

[11] Note that using the median over all simulations of the median distance over all possible paths hides the variation in distance across simulations under the null hypothesis and across nodes, and thus favors the null hypothesis of equal distance in the data and in the simulations, which is one

is that they only consider the network of boards and directors who sit on at least two boards, i.e. interlockers.

We have argued in Section 2.1 that the Q-statistic is problematic. To illustrate this point, we compute a Q-statistic, based on this counterfactual, and we compare it to the outcome of the small world test. We follow the common practice in the literature and (arbitrarily) choose a cut-off value of 2. For board projections, whenever we have small worlds using our test, the Q-statistic is nearly always larger than 2, in 100% (87%) of the cases when we simulate from the whole network (giant component). However this result is entirely due to chance, as the Q-statistic is also above the cut-off, when we do not have small worlds in 97% (74%) of the cases when we simulate from the whole network (giant component). For director projections, the Q-statistics are nearly always below the cut-off value, regardless of whether our test finds small worlds or not. Moreover, this result is robust to changing the cut-off value. This shows that the Q-statistic is not a reliable measure of the presence of small worlds.

4.3 Degree distribution of boards

The results of this second counterfactual are displayed in Panel B of Table 2. The first columns shows results of the simulations from the whole network. The results for board projections are very similar to the ones of the previous counterfactual: the simulated largest components are larger than the observed ones; we have excess clustering in almost all countries; the number of countries with observed distances equal to the simulated ones increases from 17 to 20, as does the number of countries with small worlds (from 19 to 23 countries).

However the picture changes when we consider directors instead of boards. There are now only 19 countries with excess clustering (against 35 before) and 17 countries whose observed distance is compatible with the simulation (against 16 before). This

element of the small world hypothesis.

results in a small increase to 19 countries with small worlds, but also in an increase from 11 to 15 countries compatible with random affiliation.

A comparison between simulations for the whole network and simulations based on the giant component reveals similar patterns for clustering and distance as in the first counterfactual. Clustering hardly changes for the board projections, but it increases dramatically for directors. There is a reduction in the number of countries whose distance is consistent with simulation, but this phenomenon is stronger than before, with reductions from 20 (17) to 12 (12) in the board (director) projection. This has implications for the number of countries with small world, which decreases from 23 (19) to 14 (14) in board (director) projection. In terms of the number of countries consistent with random affiliation, there is a slight increase for the board projection, from 3 to 4, but a dramatic decrease in the director projections, from 15 to 2.

4.4 Degree distributions of boards and directors

The results of this last counterfactual, displayed in Panel C of Table 2, are very different from the previous ones. We first comment the results for the giant component when we simulate from the whole network. As for the previous counterfactuals, the size of the simulated giant components is larger than the observed one, in all but 14 (12) countries, for which equality cannot be rejected, for board (director) projections. Compared to the previous counterfactuals, the distribution of simulated distances shifts downwards, while clustering shifts upwards. This implies, that there is much less excess clustering than before. There are now only 26 (3) countries with excess clustering in board (director) projections, compared to 49 (35) for the first counterfactual. The number of countries whose observed distance is not significantly different from the simulated ones increases from 17 (16) to 21 (24) for board (director) projections, but there is an inversion in the sign of the rejections. The number of small world countries is now 21 (5) for board (directors), but with

16 (7) countries for which we cannot reject random affiliation, the evidence in favor of small worlds is rather bleak.

The same shifts in the distributions of distance and clustering are at work in the simulations based on the giant component only, when compared with previous counterfactuals, e.g. Panel A. The number of countries with excess clustering is now only 34 (39), which is more than the results in Panel C from the whole network, but less than previous counterfactuals, where it was around 49. The number of countries with equal distance reduces to 8 for both boards and directors, which is the smallest of all counterfactuals that we consider. As a result, the number of small world countries is down to 9 (8), with 9 (7) countries that are consistent with random affiliation for boards (directors) projection. Thus there is no evidence of small worlds with this counterfactual, which imposes both the degree distributions of boards and directors.

So far we have not commented the results about average degree. It is instructive to compare the results of average degree between all three counterfactuals. As shown mathematically in Newman et al. (2001) and Newman, Strogatz & Watts (2002), average degree and distance both depend on the degree distributions. More specifically, the average degree of boards (directors) depends on the degree distribution of directors (boards). For instance, it is the length of the right tail of the degree distribution of the board, as measured for instance by skewness, which leads to higher average degree for the directors. A large board with, say, 10 seats will lead to 10 directors that each have a least degree 9, since they all know each other. This is confirmed in Table 2. In Panel B, when the degree distribution of the board is equal to that of the data, the results of the average degree for the boards are identical to those of Panel A, since in both cases the degree distribution of the directors is random. Instead, for directors, the results of average degree are identical between Panels B and C, since in both cases the degree distribution of the boards are equal to that of the data. For both directors and boards, the degree distribution of the data is more skewed than the random distributions that we generate in Panel A, which produces higher average degree. When the degree distributions in the simulations are very different from the ones in the data, it is the degree distribution that explains most of the results with average degree. To some degree, our explanation for average degree also holds true for distance. Distance is also affected by the degree distributions, but in a somewhat more intricate way, since it depends not only on average degrees (first neighbors) but also on the average number of second neighbors. Of course, when simulations are carried out from the entire network, there is also an effect of the size of the giant component which affects average degree and leads to lower distance.

We also report the board and director projection results of average degree, the clustering coefficient, and average distance for every country when simulating from the giant component in Tables 3 and 4. For each measure we report the actual values, the average of the counterfactual as well as the z-scores, with stars when we cannot reject the null that the measure is compatible with its empirical distribution based on a nonparametric test. These tables illustrate the fact that even though actual and simulated average values sometimes look quite close, they appear very significantly different, when their variability under the null is taken into account. The z-scores are often huge and in some cases, actual values are even entirely outside the range of the simulations. The tables also show systematically lower average degree, as well as higher distance and clustering in the data than in the simulated values for both projections. Finally, Table 5 shows the p-values for the random affiliation and small world tests. In the very large majority of cases we very strongly reject both hypotheses, as can be seen from the p-values.

5 Discussion

5.1 Implications for boards of directors networks

Taken together, the results of our three counterfactuals offer at best fairly weak evidence for the prevalence of small worlds, either for boards or directors projections. Focusing on board projections, the largest set of small world countries obtains when we simulate from the entire network and fix the degree distribution of the board. In that case, there are 20 countries with small worlds (23, three of which are ruled out since they are also compatible with random affiliation). If instead we consider directors projections, there are at most 15 small world countries (17, two of which are compatible with random affiliation), when we simulate from the largest component and keep density fixed. Thus in any case, there are never more than half the countries that are compatible with the small world hypothesis.

Moreover, for 18 countries which comprise some of the major economies in the world (e.g. the U.S., Japan, Germany, France, Italy, Brazil, Russia, India, Hong Kong, Malaysia, Thailand, Philippines), we reject the small world hypothesis in boards projections for all three counterfactuals, regardless of wether we simulate from the entire network or just from the giant component. This runs contrary to some of the findings in the literature, in particular about the U.S. (Davis et al. 2003, Robins & Alexander 2004).

The list of small world countries, for which we cannot reject the null of small worlds but which are incompatible with random affiliation, reveals interesting patterns. Although, as shown in Table 2, the total number of small world countries can be quite close when comparing different counterfactuals, the countries that actually qualify as small world can be quite different. Setting aside the last counterfactual, one can roughly consider two (partly overlapping) groups of countries. There are countries that are small worlds for simulations based on the entire network, and those that are small worlds when simulations are based on the giant component only. The first group is composed of a core of 16 countries that are small worlds for the two counterfactuals, and 4 additional countries that are only small worlds when we fix the board distributions in the simulation.

There are only 10 core countries in the second group, with an additional 5 for the first counterfactual. Moreover, there are only 3 common countries to the core of both groups. This fact is explained by the switch in the distribution of distance when we move from simulating from the whole network to the largest component. Basically the distance in the simulation for all countries decreases as we move from simulating from the whole network to the largest component. This implies that the countries where the distance is similar to the one simulated when we consider the whole network, become countries where the distance is larger than the simulated from the largest component.

When instead we consider directors projections, there are 19 countries, that also comprise some of the major economies in the world (e.g. the U.S., Canada, Australia, Germany, France, Brazil, Russia), for which we always reject the small worlds hypothesis. Table 2 shows that there are very few small world countries in directors projections when we simulate from the whole network, while there are around 15 countries in simulations from the giant component, when we focus on the two first counterfactuals. This is due to the fact that simulating from the whole network produces clustering similar to the one observed in the data, which means that there are many countries that are compatible with random affiliation. Instead we obtain excess clustering when we simulate from the giant component. Moreover, for director projections one can only identify a group of 9 countries that appear to be small worlds in the first two counterfactual when we simulate from the giant component. As in board projections, the last counterfactual, which fixes degree distributions for both boards and directors, produces only one small world country when simulating from the giant component, while in the case of the board projection it produces six small world countries when simulating from the whole network.

Since we do not have that many small world countries, it is natural to ask whether this is because many countries are random draws from the counterfactuals. As it turns out, this is not the case, since there are 36 (35) countries for which we always reject the hypothesis of random affiliation for board (director) projection. Moreover, in about a third of our sample, more precisely for 18 (19) countries for boards (directors) projections, we always reject the small world hypothesis and the random affiliation, indicating that their networks are neither random draws from our counterfactuals nor small worlds.

5.2 Methodological implications

The previous subsection reveals some of the difficulties with tests of the small world hypothesis. The main problem is that the outcome of the tests can depend heavily on the chosen counterfactual. Different countries appear to be small worlds with different counterfactuals. This is especially problematic when interest lies in a particular country. If different counterfactuals give different answers, then which one can be trusted? This perfectly illustrates the joint hypothesis problem that small worlds pose, which current research does not seem to have fully acknowledged so far. In a sense, tests of the null hypothesis of small worlds are like the tests of market efficiency that are being conducted in the finance literature in that they both suffer from the joint hypothesis problem (see Fama 1970, Fama 1991, Malkiel 2003). The problem about identifying departures from market efficiency is that one needs to know how to characterize the behavior of stock prices when markets are efficient. Thus one needs to specify an asset-pricing model. In other words, the problem is the choice of the proper counterfactual. When the null hypothesis of market efficiency is rejected, it is impossible to know if markets are truly inefficient, or whether it is the particular model of market efficiency that is incorrect. There is exactly the same problem with tests of the small worlds hypothesis, since small worlds are identified, following Watts & Strogatz (1998) as having similar distances but more clustering than “equivalent random graphs”. The identification of small worlds hinges on how one operationalizes the notion of “equivalent random graphs”. When the hypothesis of small worlds is rejected, it is impossible to know if there really are no small worlds or if the counterfactual was not the right one.

Whereas a priori one could think that imposing more structure on the counterfactuals necessarily reduces the gap between simulated and observed networks, this is no longer the case when the variability under the null hypothesis is taken into account in a proper test. A smaller gap in the difference of two measures can lead to more significant departures from randomness, if at the same time the variability under the null is reduced even more.

Moreover the small world hypothesis involves simulated distances that are close to the observed ones, while it requires observed clustering to be significantly larger than in simulations. Thus a counterfactual that is so close to the data that it matches observed clustering, will lead to rejecting the small world hypothesis in favor of the random affiliation hypothesis. It is important to stress that the correct counterfactual for a small world test is not necessarily the one that best fits observed data. By definition with the best-fitting counterfactual, the data could not produce excess clustering and this counterfactual would therefore necessarily lead to rejecting the small world hypothesis.

While there is no definite “right” counterfactual, we do see problematic issues with some of the counterfactuals that we use. First of all, it is hard to believe the counterfactuals based on simulating from the entire network, as most of those produce giant components whose sizes are significantly different from observed ones. Also, the two first counterfactuals are very different from the data in at least one of the degree distributions. Again, the random graphs under these two counterfactuals are unable to match important features of the data. Thus our preference goes to the last counterfactual, when simulating from the giant component only. This is also in line with the choice made by Conyon & Muldoon (2006).

For the sake of comparison, we also show in Table 7 a summary of results obtained under the (strictly speaking, mostly wrong) assumption of normally distributed network measures, with z-scores for individual measures, a Wald test for random affiliation and the original Kodde & Palm (1986) test for the small world hypothesis. The differences with the results of nonparametric tests is minimal, there is usually at most a few countries whose test outcomes are affected, which illustrates the fact that in the vast majority of countries, the rejection of small worlds is very strong, and robust to the use of a somewhat different testing methodology. Nonetheless, when interest focuses on a particular country, the nonparametric results are to be preferred, since they are correct regardless of the underlying distribution of the counterfactual measures.

5.3 Bicliques in bipartite graphs

The analysis so far has focused exclusively on the unipartite projections of the bipartite network of board membership. Instead, we now turn our attention to bicliques, which can only be seen in the bipartite network. The smallest such biclique is a $(2, 2)$-biclique, which obtains whenever two directors sit together on the two same boards. More generally, a (b, d)-biclique consists of two sets of $b \geq 2$ boards and $d \geq 2$ directors, such that all d directors sit on all b boards. As an illustration, Figure 2 provides four examples of small bicliques. A $(2, 2)$-biclique is depicted as the first motif on the left. Adding a third board and connecting it to the 2 existing directors creates a $(3, 2)$-biclique (second from the left). Observe the latter motif contains 3 partly overlapping $(2, 2)$-bicliques. Adding yet another director and connecting him to all three boards creates a $(3, 3)$-biclique (last on the right), which this time contains 9 partly overlapping $(2, 2)$-bicliques. In general, any (b, d)-biclique embeds large numbers of smaller order bicliques, i.e., (b', d')-bicliques with $b' \leq b$ and $d' \leq d$. It is easy to check, for instance, that an isolated (b, d)-biclique, i.e. one that has no overlap with other bicliques, spans $[b(b-1)/2] \times [d(d-1)/2]$ $(2, 2)$-

bicliques. Although each high order biclique contributes to the $(2,2)$-biclique count and to the number of redundant (and thus lost) edges in the projections, there is no direct mathematical link between these quantities since there is often partial overlap between the various bicliques. To sum up, a biclique of order higher than $(2,2)$ contributes greatly to the count of $(2,2)$-bicliques without necessarily involving many directors or many boards. In the sequel we call "maximal" $(b,d)-$bicliques, bicliques that are not subsets of any larger biclique.

What makes bicliques important is that when the bipartite network is projected onto one of its constituent sets of nodes, some connections are lost. A $(2,2)$-biclique for instance creates an edge between the two firms involved, but one common director would be enough for the edge to exist. Symmetrically this biclique creates an edge between the two directors involved, but common membership to only one board would be sufficient. In both projections, one edge is lost. More generally, an isolated (b,d)-biclique results in $(d-1)\times b(b-1)/2$ redundant links in the board projection, and $(b-1)\times d(d-1)/2$ redundant links in the director projection. This implies that the unipartite projections of a bipartite network containing a biclique has fewer edges, and thus lower average degree, than they would if the original bipartite network did not have any biclique. This in turn has implications on other structural properties of the network, such as average distance.

As shown in Table 2, there are significantly more $(2,2)$-bicliques in the data for all counterfactuals and types of simulations that we consider– a feature also present in the study of Robins & Alexander (2004). There are at most 3 countries whose count of $(2,2)$-bicliques is compatible with random affiliation when we simulate from the giant component with fixed degrees for boards and directors.

To get a clearer picture of the presence and implications of high order (b,d)-bicliques with $b,d\geq 2$, we identify all bicliques of any order in the data. Table 6 reports the number of $(2,2)$-bicliques, as well as the number of maximal (b,d)-bicliques in the bipartite data networks. The distinction is important since, as a comparison of the first and second columns of Table 6 illustrates. The count of $(2,2)$-bicliques, which includes the ones that are subsets of higher order bicliques, vastly exceeds the count of maximal $(2,2)$-bicliques. The difference, which is sometimes close to an order of magnitude, informs about the considerable extent of biclique overlap in some countries. A second pattern visible in the data is that there are many high order bicliques in large economies, although some large economies also have few of them (e.g. Australia, Japan or the U.K.). Medium-sized countries have sometimes spectacular biclique counts as well, see e.g. Sri Lanka or the Philippines. Finally small countries always have very few motifs of order higher than $(2,2)$.

The number of redundant edges in the projections has an impact on average degree and on most other measures we have considered so far. For now we focus attention on the last counterfactual (giant component and preserved degree distributions) in order to avoid systematic differences between data and simulation outcomes that would confound the effects which we try to isolate.

First, for given degree distributions in the bipartite network, average degree in the unipartite projections depends exclusively on the number of redundant edges, of which the count of $(2,2)$-bicliques is a good proxy. Indeed, multiple common directors or multiple memberships are not considered in unipartite projections, where multiple edges are projected onto a single one. Since it is common in the data but rare in the counterfactuals that two firms share more than one director, the unipartite projections of counterfactual networks will have fewer edges that are redundant relative to the unipartite projections of the original bipartite data. Therefore the counterfactual projections will contain more edges than the data projections, and thus display higher average degree (every redundant edge lost in the projection reduces the degree sum by 2). We observe this effect at the aggregate level in panel C of Table 2 and at the country level in Tables 3 and 4 for boards and directors. Degree distributions in the bipartite counterfactual are equal to the ones in the actual data, but as the $(2,2)$-biclique count in the data is always greater than or equal to the $(2,2)$-biclique count in the counterfactual (strictly greater in 48 countries, equality cannot be rejected in 5 countries), the average degree of the unipartite projections is larger in data than in simulation outcomes.

Second, all else equal, having more redundant edges in the projections and thus fewer actual edges on average increases distances, since these (lost) edges would, if reintegrated, contribute to shortening at least some of the shortest paths, provided that no multiple links are allowed. This explains why average distance in the actual data is mostly above average distance in the counterfactuals. This finding is consistent with the results of Zhang & Zhang (2009), who document that the diameter of empirical networks in many different fields is smaller than the one of random counterfactuals, because empirical networks tend to display more modularity, which is a property related to the presence of cliques. However, the magnitude of this effect cannot be quantified a priori, as it depends on the degree of the nodes that are involved in repeated edges.

We provide graphical evidence of the link between redundant edges and the z-scores of distance and clustering in board projections in Figure 3. The figure shows that the gap between simulated and actual data (measured in number of standard deviations) is larger in countries with more redundant edges. We further investigate this relation by regressing the z-scores on the count of redundant edges controlling for network size. We use the log of all variables, and in the case of z-scores, we take the absolute value because of the very limited number of small negative values that appear with few redundant edges (bottom left of the plots in Figure 3). Results in Table 8 show that for both projections, when controlling for size, the elasticity of the z-score with respect to the number of redundant edges is close to 1, which means that a 1% increase in the number of redundant edges increases the z-scores of distance and clustering by 1%. Besides, the effect of size is always negative and significant, and the fit of the regressions is quite high, with R^2s of 84% and 88% for distance, and 70% and 61% for clustering, respectively in the board and director projections. This confirms that in our 53 countries, redundant edges are systematically associated with larger z-scores for clustering and distance, i.e., larger deviations from the data, and this is true even when we control for the size of the network. While in the case of clustering, larger z-scores imply increased excess clustering in the data and thus increased support for the small world hypothesis, larger z-scores for distance imply an increased departure from the assumption of 'random' shortest paths and thus rejection of the small world hypothesis in most countries.

To summarize, the presence of sometimes numerous bicliques in the bipartite data network leads to many redundant edges, the loss of which lowers the density that we observe in the data projections. In general lower density implies longer distances, and thus it is mostly because of this significantly larger distance in the data projections compared to the counterfactual projections that we reject the small world hypothesis in most countries. It is worth emphasizing that we mostly reject the small world hypothesis not because the data is compatible with random affiliation, but simply because distances are too large. It now appears that this is due to a large extent to the existence of highly dense subsets of the graph — small "clubs" of directors who sit jointly on several boards — who cannot form just by chance but rather point at the existence of additional, non-random social processes. This raises the question whether in bipartite graphs, it is not precisely these very localized motifs which should be the center of attention, as they affect the global structure of the network and create important departures from 'comparable' random graphs. From a managerial perspective, bicliques offer support to a view of corporate governance as embedded in social structures such as formulated in Davis (1996) for instance. Davis (pp 157-158) argues that rather than pecuniary incentives, "...directors are more likely to respond to concerns about honor, obligations, and notions of appropriateness...", also stating that "New directors are commonly recruited through friends or acquaintances that are already on the board, often the

CEO." (p. 157) and that "...to the extent that incentives or shareholder pressures are to be effective, it is through their impact on the social relations among peers." (p. 158). The multiplicity of high-order bicliques, in which the same people meet in the same boards, inviting members of one board to join another one, and possibly reciprocating received invitations, offers strong support to the hypothesis of embeddedness.

6 Conclusion

Small worlds seem ubiquitous in many different fields, from social networks to biological, and even technological networks. While it might be that small worlds are universal as claimed, one should also recognize that both the notion of a small world as defined in the seminal paper by Watts & Strogatz (1998) and the identification procedures used in papers analyzing empirical networks lack some elements of rigor and precision. In this paper we aim to fill this gap and provide a rigorous testing method that relies on simulating appropriately defined counterfactual networks. We apply our testing methodology to the structure of board membership networks in a large number of countries.

More specifically, we offer a principled discussion of counterfactual design and introduce a formal explicitly joint nonparametric test of the small world hypothesis, based on distance and clustering. Our method could more generally be used in correctly sized joint tests on a set of unipartite or bipartite network measures, to check the overall adequacy of a sample of counterfactuals with a given observation, while accounting for the possible correlations between the different measures.

Our test rejects the small world hypothesis for corporate board networks for all counterfactuals, and for most countries, especially the largest ones. When we simulate from the giant component for the counterfactual that imposes the degree distributions of boards and directors, we reject the small world hypothesis for board projections in all but one country. Our results highlight the importance of controlling for the variability and correlation of the theoretical measures, a feature which for instance the small world quotient, a widely used identification "statistic", fails to do.

Our findings hold across a wide variety of economic contexts, institutional frameworks, capital markets and traditions of corporate governance. In accordance with earlier research, we generally find excess clustering, even though we explicitly operate on bipartite counterfactuals. However, average distance is almost always greater in the data than in the counterfactual, causing the joint test to fail. Similarly, Zhang & Zhang (2009) also find distances that are greater than expected in a number of empirical networks, and they attribute this feature to the presence of modularity. Their explanation is closely related to our conjecture that the discrepancy between the data and counterfactual networks is due to the presence of bicliques in the bipartite networks, which lead to lost edges in the unipartite projections. Our evidence suggests that the presence of lost edges contributes strongly to the patterns found in the data, particularly in terms of distance in the unipartite projections. Having small groups of 2, 3 or 4 directors who jointly sit on 2, 3 or 4 boards creates a multiplicity of relationships that randomness in board affiliation is never able to reproduce, and explains much of the departure from the small world hypothesis. This feature, that overall structure is not impacted so much by directors with unusually large degree but rather by the accumulation of certain directors in certain boards, is already suggested in Robins & Alexander (2004). It is likely that these cliques of directors, who participate in several boards together, obey logics of social embeddedness, as advocated by Davis (1996), forming through chains of referrals and reciprocated invitations. Whatever their origin might be, these high-order motifs, which have been shown to be relevant in other disciplines (see for instance Alon 2007) have the remarkable feature of exerting a strong influence on the global properties of the overall network structure while being very localized phenomena.

In this paper we find that when put to a strict statistical test, only few board or director projections qualify as small worlds. Our findings run contrary to most empirical work in the small worlds literature. One possibility is that our results are due to the strong characterization we adopt. Unlike previous literature, we combine the demands of excess-relative-to-random clustering and quasi-random distance in a statistical test. Nonetheless, our results call for some caution in the identification of small worlds. Going back to the U.S. corporate network, whose giant component encompasses 30223 directors, we find an average distance in the data of 6.71 (parenthetically, a beautiful instance of 'six-degrees-of-separation'), whereas the average distance in the last counterfactual is 5.90. Even though, when corrected for variability in the counterfactual, the difference between the two, measured in terms of z-scores becomes as high as 52.80, it is tempting to conclude that a world of thirty thousand network actors scattered over an entire continent with an average distance between two actors of 6.71 is a fairly small one indeed, if not in a statistical sense. However, in that particular case, as in many of our tests, while observed and simulated network measures seem numerically quite close, the observed value is actually far outside the range of possible values obtained in 1000 simulations of the counterfactual. This points to the misunderstanding that consists in comparing empirical and simulated network measures without considering their range and variability. Alternatively, our rejection of small worlds may also point to an issue in the definition of small worlds. One might be able to gain some insight into this question by applying our small world test to other unipartite and bipartite networks that have been identified as small worlds in the extant literature.

References

Adams, R. B., Hermalin, B. E. & Weisbach, M. S. (2010), 'The role of boards of directors in corporate governance: A conceptual framework and survey', *Journal of Economic Literature* **48**(1), 58–107.

Alon, U. (2007), 'Network motifs: Theory and experimental approaches', *Nature Reviews Genetics* **8**(3), 450–461.

Batagelj, V. & Mrvar, A. (2000), 'Some analyses of Erdős collaboration graphs', *Social Networks* **22**(2), 173–186.

Baum, J., Shipilov, A. V. & Rowley, T. J. (2003), 'Where do small worlds come from?', *Industrial and Corporate Change* **12**, 697–725.

Bouwman, C. (2011), 'Corporate governance propagation through overlapping directors', *Review of Financial Studies* **24**(7), 2358–2394.

Conyon, M. J. & Muldoon, M. R. (2006), 'The small world of corporate boards', *Journal of Business Finance and Accounting* **11**, 1321–1343.

Conyon, M. J. & Muldoon, M. R. (2008), Ownership and control: A small world analysis, *in* J. Baum & T. Rowley, eds, 'Network Strategy: Advances in Strategic Management', Oxford, pp. 31–66.

Davis, G. (1991), 'Agents without principles? The spread of the poison pill through the intercorporate network', *Administrative Science Quarterly* **36**, 583–613.

Davis, G. (1996), 'The significance of board interlocks for corporate governance', *Corporate Governance* **4**, 154–159.

Davis, G. & Greve, H. R. (1997), 'Corporate elite networks and governance changes in the 1980s', *American Journal of Sociology* **103**, 1–37.

Davis, G., Yoo, M. & Baker, W. (2003), 'The small world of the corporate elite: 1982-2001', *Strategic Organization* **1**(3), 301–326.

De Castro, R. & Grossman, J. W. (1999), 'Famous trails to Paul Erdős', *The Mathematical Intelligencer* **21**(3), 51–53.

Erdős & Rényi (1959), 'On random graphs. I', *Publicationes Mathematicae (Debrecen)* **6**, 290–297.

Fama, E. (1970), 'Efficient capital markets: A review of theory and empirical work', *Journal of Finance* **25**(2), 383–417.

Fama, E. (1991), 'Efficient capital markets II', *Journal of Finance* **46**(5), 1571–1617.

Fleming, L., King, C. & Juda, A. (2007), 'Small worlds and regional innovation', *Organization Science* **18**(6), 938954.

Goyal, S., van der Leij, M. & Moraga-González, J.-L. (2006), 'Economics: An emerging small world', *Journal of Political Economy* **114**(2), 403–412.

Grossman, J. W. & Ion, P. D. F. (1995), 'On a portion of the well-known collaboration graph', *Congressus Numerantium* **108**, 129–132.

Gulati, R., Sytch, M. & Tatarynowicz, A. (2012), 'The rise and fall of small worlds: Exploring the dynamics of social culture', *Organization Science* **23**(2), 449–471.

Hallock, K. (1997), 'Reciprocally interlocking board of directors and executive compensation', *Journal of Financial and Quantitative Analysis* **32**, 331–344.

Haunschild, P. (1993), 'Interorganizational imitation: The impact of interlocks on corporate acquisition activity', *Administrative Science Quarterly* **38**, 564–592.

Hayfield, T. & Racine, J. S. (2008), 'Nonparametric econometrics: The np package', *Journal of Statistical Software* **27**(5).

Hermalin, B. & Weisbach, M. (2003), 'Boards of directors as an endogenously determined institution: A survey of the economic literature', *Economic Policy Review of the Federal Reserve Bank of New York* **9**, 7–26.

Hillman, A. J., Cannella, A. A. J. & Paetzold, R. L. (2000), 'The resource dependence role of corporate directors: Strategic adaptation of board composition in response to environmental change', *Journal of Management Studies* **37**(2), 1–10.

Humphries, M. D., Guerney, K. & Prescott, T. J. (2006), 'The brainstem reticular formation is a small world, not scale-free, network', *Proceedings of the Royal Society B: Biological Sciences* **273**(1585), 503–511.

Hyndman, R. J. (1996), 'Computing and graphing highest density regions', *The American Statistician* **50**(2), 120–126.

Kodde, D. A. & Palm, F. C. (1986), 'Wald criteria for jointly testing equality and inequality restrictions', *Econometrica* **54**(5), 1243–1248.

Kogut, B. & Walker, G. (2001), 'The small world of Germany and the durability of national networks', *American Sociological Review* **66**, 317–335.

Malkiel, B. G. (2003), 'The efficient market hypothesis and its critics', *Journal of Economic Perspectives* **17**(1), 5982.

Milgram (1967), 'The small world problem', *Psychology Today* **1**(1), 60–67.

Milo, R., Kashtan, R., Itzkovitz, S., Newman, M. E. J. & Alon, D. (2004), 'On the uniform generation of random graphs with prescribed degree sequences', *arXiv:cond-mat/0312028v2* .

Mizruchi, M. S. & Stearns, L. (1994), 'A longitudinal study of borrowing by large American corporations', *Administrative Science Quarterly* **39**, 118–140.

Newman, M. (2001*a*), 'Scientific collaboration networks. I. Network construction and fundamental results', *Physical Review E* **64**, 016131.

Newman, M. (2001*b*), 'Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality', *Physical Review E* **64**, 016132.

Newman, M. (2001*c*), 'The structure of scientific collaboration networks', *PNAS* **98**(2), 404–409.

Newman, M. E. J. (2004), Who is the best connected scientist? A study of scientific coauthorship networks, *in* E. Ben-Naim, H. Frauenfelder & Z. Toroczkai, eds, 'Complex Networks', Springer, pp. 337–370.

Newman, M., Strogatz, S. & Watts, D. (2001), 'Random graphs with arbitrary degree distribution and their applications', *Physical Review E* **64**(026118).

Newman, M., Strogatz, S. & Watts, D. (2002), 'Random graph models for social networks', *PNAS* **99**, 2566–2572.

O'Reilly, C. I., Main, B. & Crystal, G. (1988), 'CEO compensation as tournament and social comparison: A tale of two theories', *Administrative Science Quarterly* **33**, 257–274.

Palmer, D., Devereaux, J. & Zhou, X. (1993), 'Late adoption of the multidivisional form by large U.S. corporations: A further exploration of the institutional, political, and economic accounts', *Administrative Science Quarterly* **37**(1), 100–131.

Phelps, C., Heidl, R. & Wadhwa, A. (2012), 'Knowledge, networks, and knowledge networks: A review and research agenda', *Journal of Management* **38**(4), 1115–1166.

Robins, G. & Alexander, M. (2004), 'Small worlds among interlocking directors: Network structure and distance in bipartite graphs', *Computational and Mathematical Organization Theory* **10**, 69–94.

Rosenkopf, L. & Padula, G. (2008), 'Investigating the microstructure of network evolution: Alliance formation in the mobile communications industry', *Organization Science* **19**(5), 669–687.

Shipilov, A. V., Greve, H. R. & Rowley, T. J. (2010), 'When do interlocks matter? Institutional logics and the diffusion of multiple corporate governance practices', *Academy of Management Journal* **53**(4), 846–864.

Stock, J. H. & Watson, M. W. (2010), *Introduction to Econometrics*, Addison-Wesley Series in Economics, Third Edition, Prentice Hall, Pearson Education.

Uzzi, B., Amaral, L. A. & Reed-Tsochas, F. (2007), 'Small-world networks and management science research: A review', *European Management Review* **4**, 77–91.

Uzzi, B. & Spiro, J. (2005), 'Collaboration and creativity: The small world problem', *American Journal of Sociology* **111**(2), 457–504.

Watts, D. J. (1999), 'Networks, dynamics, and the small-world phenomenon', *American Journal of Sociology* **105**(2), 493–527.

Watts, D. J. & Strogatz, S. H. (1998), 'Collective dynamics of 'small-world' network', *Nature* **393**, 440–442.

Zhang, Z. & Zhang, J. (2009), 'A big world inside small-world networks', *PLoS ONE* **4**(1), e5686.

Table 1: Descriptive statistics

Country	Whole network				Giant Component			
	Number of Boards	Number of Directors	Board Size	Seats per director	Number of Boards	Number of Directors	Board Size	Seats per director
Argentina	54	466	9.76	1.13	24	265	13.58	1.23
Australia	1741	6988	5.05	1.26	1108	4371	5.35	1.36
Austria	86	650	8.52	1.13	43	378	10.42	1.19
Bahrain	36	326	10.06	1.11	20	178	10.45	1.17
Bangladesh	104	1017	10.93	1.12	57	627	12.96	1.18
Belgium	122	1002	9.32	1.13	78	732	11.05	1.18
Brazil	230	1507	7.87	1.20	153	1045	8.76	1.28
Canada	3596	15763	6.06	1.38	2817	12485	6.51	1.47
Chile	121	814	7.96	1.18	86	580	8.44	1.25
China	1627	11728	7.69	1.07	409	3398	9.53	1.15
Cyprus	104	633	6.99	1.15	21	151	9.14	1.27
Denmark	142	998	7.59	1.08	56	437	8.93	1.14
Finland	120	873	8.43	1.16	96	741	9.13	1.18
France	564	3688	7.62	1.16	226	2181	12.13	1.26
Germany	735	4893	7.38	1.11	241	2718	13.05	1.16
Greece	200	1543	8.33	1.08	38	388	11.84	1.16
Hong Kong	1167	8345	8.73	1.22	856	6207	9.35	1.29
India	3187	19298	7.33	1.21	1745	11015	8.45	1.34
Indonesia	294	1153	4.48	1.14	82	355	5.66	1.31
Ireland	83	713	9.18	1.07	26	283	12.08	1.11
Israel	347	1620	5.71	1.22	194	1119	7.55	1.31
Italy	256	3292	15.28	1.19	205	2801	16.67	1.22
Jamaica	35	258	8.83	1.20	24	186	9.83	1.27
Japan	2396	25415	11.22	1.06	879	10891	13.74	1.11
Jordan	146	1004	7.42	1.08	23	213	10.57	1.14
Kenya	43	332	8.81	1.14	30	238	9.40	1.18
Kuwait	149	914	6.73	1.10	31	196	7.35	1.16
Malaysia	968	5791	7.49	1.25	739	4329	7.73	1.32
Mauritius	30	213	9.73	1.37	27	192	10.04	1.41
Mexico	90	918	12.56	1.23	65	664	13.43	1.31
Netherlands	153	948	6.85	1.11	64	450	8.34	1.19
New Zealand	136	745	6.24	1.14	63	349	6.81	1.23
Nigeria	156	1291	9.23	1.12	77	644	9.78	1.17
Norway	203	1828	10.17	1.13	137	1471	12.39	1.15
Oman	99	621	7.24	1.15	60	390	7.88	1.21
Pakistan	328	1976	7.66	1.27	202	1174	8.10	1.39
Philippines	223	1538	9.71	1.41	183	1223	10.09	1.51
Poland	275	1434	5.75	1.10	32	187	6.94	1.19
Portugal	42	530	13.67	1.08	27	385	15.85	1.11
Qatar	42	298	7.88	1.11	25	188	8.72	1.16
Russia	199	1841	10.52	1.14	105	1053	12.26	1.22
Saudi Arabia	116	886	8.10	1.06	31	253	9.32	1.14
Singapore	594	3466	7.26	1.24	474	2711	7.47	1.31
South Africa	338	2625	9.35	1.20	257	2098	10.20	1.25
Spain	136	1422	11.63	1.11	88	1022	13.35	1.15
Sri Lanka	219	1098	7.79	1.55	192	933	7.99	1.65
Sweden	373	2211	7.22	1.22	262	1721	8.33	1.27
Switzerland	274	1732	7.10	1.12	142	953	8.06	1.20
Taiwan	479	2578	5.70	1.06	60	448	8.85	1.19
Thailand	499	4467	10.63	1.19	374	3380	11.23	1.24
Turkey	151	956	7.38	1.17	31	192	8.90	1.44
United Kingdom	1521	8809	6.63	1.14	781	5029	7.80	1.21
United States	9249	54764	6.95	1.17	4375	30223	8.88	1.29

This table shows descriptive statistics for our sample of 53 countries, ranked according to the number of firms in the largest component. The table displays the number of boards and directors, as well as average degree for boards (Board size) and directors (Seats per director), for all listed firms (Network) as well as the largest connected component (Connected Network).

Table 2: Summary of nonparametric tests for all simulations

Panel A: Holding density fixed

	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	37	14	2	35	15	3						
Degree	0	6	47	1	9	43	8	17	28	9	17	27
Clustering	0	4	49	2	16	35	0	3	50	0	4	49
Distance	32	17	4	33	16	4	4	17	32	7	15	31
Small world		19	34		16	37		18	35		17	36
Random affiliation		3	50		11	42		3	50		2	51
(2,2)-bicliques	0	0	53				0	1	52			

Panel B: Holding degree distribution of boards fixed

	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	37	16	0	39	14	0						
Degree	0	7	46	3	18	32	7	18	28	50	3	0
Clustering	0	3	50	10	24	19	0	4	49	0	5	48
Distance	29	20	4	29	17	7	4	12	37	3	12	38
Small world		23	30		19	34		14	39		14	39
Random affiliation		3	50		15	38		4	49		2	51
(2,2)-bicliques	0	0	53				0	2	51			

Panel C: Holding degree distribution of boards and directors fixed

	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	39	14	0	41	12	0						
Degree	16	26	11	3	19	31	49	4	0	48	5	0
Clustering	5	22	26	41	9	3	1	18	34	5	9	39
Distance	9	21	23	10	24	19	0	8	45	0	8	45
Small world		21	32		5	48		9	44		8	45
Random affiliation		16	37		7	46		9	44		7	46
(2,2)-bicliques	0	2	51				0	3	50			

This table shows summary statistics for nonparametric tests performed when the density of the data is held fixed in the simulations (Panel A), when imposing on the simulations the degrees distribution of the boards (Panel B), and when both the degree distribution of the boards and directors is imposed on the counterfactual (Panel C). The table shows the number of countries (out of a total of 53 countries with large enough networks) that have simulated measures that are smaller ("<"), not significantly different ("=") or greater (">") than the actual observed measures. Equality corresponds to a bilateral test at 10%, while inequalities correspond to unilateral tests at 5%. Simulations are performed from the entire network in the left half of the table and with data from the giant component only in the right half of the table. Largest component compares the size of actual with simulated giant components. Degree, distance and clustering are as shown in Appendix A. Small world refers to the nonparametric test of equal distance and excess clustering described in Appendix D.2, while random affiliation refers to the nonparametric joint test for distance and clustering described in Appendix C.2. $(2,2)$-bicliques refer to the $(2,2)$-biclique count. Tests of individual measures are based on the quantiles of the distribution of the counterfactual measures.

Table 3: Board projections of random graphs, holding degree distribution of boards and directors fixed (simulating from the giant component)

Country	Average degree			Clustering coefficient			Average distance		
	Obs.	Sim.	z-score	Obs.	Sim.	z-score	Obs.	Sim.	z-score
Argentina	3.92	6.44	-9.37	0.71	0.48	4.67	3.09	1.92	17.03
Australia	3.96	4.41	-130.94	0.51	0.40	16.33	8.39	5.62	51.18
Austria	3.21	4.07	-9.38	0.35	0.34	0.28***	3.84	2.93	7.89
Bahrain	3.90	3.99	0.55***	0.46	0.49	0.51***	2.62	2.30	2.20**
Bangladesh	3.02	4.81	-20.50	0.47	0.28	5.38	4.39	2.82	20.72
Belgium	3.95	4.55	-10.46	0.40	0.34	1.98**	3.70	3.21	5.22
Brazil	4.92	6.55	-39.74	0.53	0.37	8.75	4.17	3.11	24.38
Canada	7.06	8.29	-372.71	0.47	0.37	27.64	5.27	4.24	70.99
Chile	3.65	5.05	-25.32	0.50	0.36	5.36	5.05	3.15	23.69
China	3.09	3.42	-50.90	0.44	0.35	7.16	10.96	6.16	31.11
Cyprus	2.67	4.61	-10.33	0.56	0.42	2.40*	3.64	2.13	17.65
Denmark	2.93	3.03	-2.73*	0.48	0.48	0.02***	4.62	4.46	0.41***
Finland	3.23	3.38	-3.81	0.28	0.29	0.38***	4.65	3.96	4.22
France	6.12	7.18	-27.29	0.35	0.28	5.26	3.91	3.08	30.51
Germany	4.29	4.73	-12.39	0.33	0.37	-1.75**	4.55	3.57	14.65
Greece	3.26	4.09	-7.87	0.50	0.39	2.63*	3.74	2.79	8.42
Hong Kong	6.28	7.86	-175.29	0.49	0.34	22.08	5.01	3.77	67.79
India	7.24	8.45	-259.78	0.45	0.35	22.12	4.85	4.02	65.60
Indonesia	2.98	3.56	-16.05	0.43	0.36	2.37*	5.67	4.03	9.73
Ireland	2.23	2.46	-3.04*	0.24	0.16	1.21***	4.14	3.67	1.41***
Israel	4.95	6.21	-39.39	0.55	0.44	7.09	4.19	3.36	16.51
Italy	6.62	8.40	-40.76	0.31	0.21	9.14	3.37	2.81	31.97
Jamaica	4.17	5.14	-4.93	0.50	0.41	1.64***	2.36	2.13	2.86*
Japan	3.20	3.56	-101.22	0.33	0.27	7.13	6.77	6.19	7.65
Jordan	2.52	2.82	-3.07*	0.37	0.26	1.54***	4.31	3.10	4.79
Kenya	5.20	5.41	-1.63***	0.52	0.56	0.80***	2.66	2.44	1.56***
Kuwait	2.13	2.23	-2.95**	0.22	0.19	0.61***	5.09	4.91	0.31***
Malaysia	5.25	6.40	-140.82	0.45	0.34	13.95	5.33	4.09	53.24
Mauritius	5.93	8.30	-8.27	0.56	0.51	1.41***	2.29	1.80	10.25
Mexico	7.23	10.09	-18.06	0.50	0.37	6.95	2.77	2.12	21.55
Netherlands	3.13	3.33	-4.31	0.38	0.41	0.62***	4.87	3.80	5.56
New Zealand	2.79	3.19	-10.16	0.37	0.30	2.16*	5.36	4.14	5.34
Nigeria	3.12	3.62	-12.22	0.39	0.31	2.73	4.44	3.79	4.53
Norway	3.84	4.08	-3.58	0.37	0.44	-2.43*	4.11	3.33	7.51
Oman	2.87	3.43	-11.34	0.44	0.29	4.40	6.02	3.75	13.74
Pakistan	4.55	7.68	-86.42	0.62	0.30	26.64	5.05	3	86.64
Philippines	8.40	13.20	-64.95	0.51	0.30	22.49	3.28	2.38	65.21
Poland	2.75	2.89	-2.76*	0.52	0.42	2.12*	4.51	3.93	1.42***
Portugal	2.81	3.43	-4.81	0.32	0.30	0.35***	3.11	2.74	2.33*
Qatar	3.12	3.17	0.58***	0.47	0.43	0.53***	2.88	3.13	0.88***
Russia	4.51	6.48	-30.85	0.52	0.30	10.39	4.40	2.83	37.27
Saudi Arabia	2.13	2.46	-5.84	0.24	0.18	1.16***	4.89	4.12	1.86***
Singapore	5.80	6.56	-58.32	0.47	0.41	6.72	4.80	3.86	28.01
South Africa	5.86	6.62	-27.92	0.39	0.33	5.04	3.97	3.35	18.55
Spain	3.52	4.28	-15.43	0.29	0.24	2.14*	4.82	3.37	17.99
Sri Lanka	9.42	15.50	-77.26	0.64	0.38	26.08	3.41	2.31	72.19
Sweden	4.95	5.62	-30.46	0.42	0.37	3.19	4.53	3.67	18.32
Switzerland	3.41	3.56	-6.77	0.35	0.33	0.90***	4.82	4.51	2.25*
Taiwan	3.13	3.86	-12.33	0.40	0.38	0.77***	5.07	3.48	9.60
Thailand	5.50	6.92	-77.63	0.39	0.27	12.89	4.41	3.49	40.13
Turkey	5.16	8.85	-13.82	0.72	0.50	6.91	3.91	1.83	47.22
United Kingdom	3.60	3.69	-20.59	0.37	0.32	4.71	6.83	5.87	12.63
United States	5.07	5.57	-376.09	0.30	0.24	18.79	6.03	5.22	55.08

This table shows degree, clustering and distance for board projections for random affiliation graphs. For each measure, the table shows actual values, average counterparts computed on equivalent simulated random networks, as well as standardized z-scores for the difference between actual and simulated measures. Stars identify cases in which we **cannot** reject the null hypothesis that the observed network has been generated randomly with equal degree distribution. ***: test at 90%, **: test at 95% and *: test at 99%.

Table 4: Directors projections of random graphs, holding degree distribution of boards and directors fixed (simulating from the giant component)

Country	Average degree			Clustering coefficient			Average distance		
	Obs.	Sim.	z-score	Obs.	Sim.	z-score	Obs.	Sim.	z-score
Argentina	17.16	17.89	-23.91	0.94	0.92	19.86	3.56	2.62	15.72
Australia	7	7.17	-191.06	0.87	0.87	14.94	8.72	6.31	41.83
Austria	13.97	14.15	-15.75	0.94	0.93	6.02	4.51	3.52	8.71
Bahrain	13.73	13.75	0.89***	0.93	0.94	-2.50*	3.39	2.97	2.95*
Bangladesh	16.69	17.05	-40.90	0.94	0.93	16.80	4.81	3.48	18.38
Belgium	14.83	14.92	-14.32	0.94	0.93	2.42*	4.53	3.88	7.28
Brazil	11.14	11.64	-82.52	0.92	0.90	27.98	4.95	3.88	22.37
Canada	10.20	10.56	-479.69	0.86	0.85	37.92	6.03	4.92	71.91
Chile	10.44	10.71	-31.66	0.92	0.91	11.25	5.79	3.92	22.94
China	11.63	11.73	-133.68	0.95	0.94	8.64	11.49	6.88	29.10
Cyprus	10.60	11.31	-23.62	0.92	0.89	16.33	4.25	2.84	16.51
Denmark	11.58	11.59	-2.66*	0.94	0.95	-1.51***	5.23	5.08	0.37***
Finland	15.19	15.21	-4.20	0.92	0.93	-5.34	5.58	4.41	7.50
France	18.23	18.50	-65.55	0.91	0.91	4.89	4.43	3.68	30.84
Germany	33.93	34.01	-17.52	0.95	0.95	-5.42	4.41	3.83	11.03
Greece	15.02	15.15	-11.06	0.94	0.94	2.36*	4.33	3.46	7.80
Hong Kong	11.98	12.38	-315.03	0.92	0.90	55.17	5.79	4.57	60.12
India	11.16	11.46	-397.20	0.90	0.89	63.38	5.66	4.81	59.45
Indonesia	7.09	7.28	-23.28	0.89	0.88	5.87	6.21	4.66	9.04
Ireland	14.29	14.32	-3.84	0.95	0.95	0.83***	4.78	4.29	1.44***
Israel	11.49	11.75	-45.47	0.91	0.90	5.85	4.70	3.91	15.46
Italy	21.37	21.63	-77.12	0.92	0.92	19.75	4.19	3.61	34.52
Jamaica	12.23	12.51	-9.42	0.91	0.90	8.37	3.02	2.82	2.60*
Japan	16.32	16.36	-126.98	0.96	0.96	6.92	7.56	6.97	7.42
Jordan	11.50	11.55	-4.57	0.94	0.94	3.27	4.92	3.84	4.30
Kenya	10.39	10.42	-2.23**	0.94	0.94	0.13***	3.40	3.28	0.83***
Kuwait	8.06	8.08	-2.93**	0.92	0.92	2.65*	5.88	5.57	0.54***
Malaysia	9.61	9.92	-217.40	0.90	0.89	50.49	6.17	4.90	49.62
Mauritius	13.35	14.02	-13.75	0.89	0.86	11.36	2.99	2.49	11.41
Mexico	17.57	18.17	-35.69	0.91	0.89	22.47	3.53	2.88	23.56
Netherlands	11.77	11.86	-12.71	0.93	0.93	0.26***	5.37	4.32	5.53
New Zealand	8.03	8.11	-10.64	0.91	0.90	5.26	6	4.86	4.93
Nigeria	12.21	12.30	-19.24	0.94	0.93	5.36	5.06	4.47	4.06
Norway	45.45	45.49	-3.69	0.95	0.96	-5.85	4.12	3.35	11.39
Oman	8.97	9.09	-15.04	0.92	0.91	10.58	6.85	4.52	13.59
Pakistan	9.70	10.51	-129.37	0.90	0.86	90.47	5.79	3.82	69.010
Philippines	14.03	15.25	-108.81	0.88	0.84	65.36	4.13	3.15	60.97
Poland	7.74	7.77	-2.72*	0.92	0.92	1.44***	5.07	4.65	1.02***
Portugal	21.75	21.81	-5.86	0.96	0.96	0.77***	3.57	3.28	1.99**
Qatar	9.78	9.78	0.55***	0.94	0.94	0.30***	3.66	3.86	0.71***
Russia	15.92	16.28	-52.87	0.93	0.92	17.34	5.01	3.57	34.16
Saudi Arabia	10.24	10.28	-5.60	0.93	0.93	1.88*	5.65	4.82	2.02**
Singapore	9.50	9.71	-95.65	0.91	0.90	17.94	5.59	4.66	24.81
South Africa	13.54	13.72	-52.32	0.92	0.91	9.53	4.60	4.11	14.26
Spain	16.68	16.80	-28.73	0.94	0.94	4.27	5.59	4.08	19.02
Sri Lanka	11.08	12.61	-98.30	0.88	0.83	80.95	4.24	3.09	62.05
Sweden	11.45	11.56	-30.79	0.91	0.91	0.19***	5.06	4.37	14.01
Switzerland	10.19	10.22	-7.25	0.92	0.92	0.17***	5.63	5.19	3.16
Taiwan	11.92	12.03	-14.15	0.94	0.93	2.37*	5.61	4.07	9.21
Thailand	13.56	13.83	-133.25	0.92	0.91	37.61	5.20	4.32	35.41
Turkey	11.21	12.01	-17.25	0.89	0.86	14.46	4.44	2.59	37.19
United Kingdom	10.30	10.31	-20.66	0.92	0.92	-6.36	7.30	6.54	9.61
United States	12.52	12.65	-641.77	0.89	0.89	4.43	6.71	5.90	51.95

This table shows degree, clustering and distance for directors projections for random affiliation graphs. For each measure, the table shows actual values, average counterparts computed on equivalent simulated random networks, as well as standardized z-scores for the difference between actual and simulated measures. Stars identify cases in which we **cannot** reject the null hypothesis that the observed network has been generated randomly with equal degree distribution. ***: test at 90%, **: test at 95% and *: test at 99%.

Table 5: Random affiliation and small world tests (p-values of nonparametric tests), holding degree distribution of boards and directors fixed (simulating from the giant component)

Country	Small world		Random affiliation	
	Boards	Directors	Boards	Directors
Argentina	0	0	0	0
Australia	0	0	0	0
Austria	0	0	0	0
Bahrain	0.085	0.002	0.106	0
Bangladesh	0	0	0	0
Belgium	0	0	0	0
Brazil	0	0	0	0
Canada	0	0	0	0
Chile	0	0	0	0
China	0	0	0	0
Cyprus	0	0	0	0
Denmark	0.632	0.234	0.780	0.349
Finland	0	0	0	0
France	0	0	0	0
Germany	0	0	0	0
Greece	0	0	0	0
Hong Kong	0	0	0	0
India	0	0	0	0
Indonesia	0	0	0	0
Ireland	0.152	0.159	0.126	0.141
Israel	0	0	0	0
Italy	0	0	0	0
Jamaica	0.016	0.028	0.006	0.002
Japan	0	0	0	0
Jordan	0.003	0.002	0.002	0.003
Kenya	0.095	0.309	0.139	0.448
Kuwait	0.609	0.487	0.593	0.054
Malaysia	0	0	0	0
Mauritius	0	0	0	0
Mexico	0	0	0	0
Netherlands	0.001	0	0.001	0
New Zealand	0	0.001	0	0
Nigeria	0.001	0.012	0.001	0
Norway	0	0	0	0
Oman	0	0	0	0
Pakistan	0	0	0	0
Philippines	0	0	0	0
Poland	0.137	0.227	0.043	0.173
Portugal	0.036	0.060	0.063	0.065
Qatar	0.658	0.911	0.644	0.801
Russia	0	0	0	0
Saudi Arabia	0.079	0.121	0.058	0.029
Singapore	0	0	0	0
South Africa	0	0	0	0
Spain	0	0	0	0
Sri Lanka	0	0	0	0
Sweden	0	0	0	0
Switzerland	0.052	0.014	0.075	0.025
Taiwan	0	0	0	0
Thailand	0	0	0	0
Turkey	0	0	0	0
United Kingdom	0	0	0	0
United States	0	0	0	0

This table shows the p-values of the nonparametric tests of the hypotheses of small worlds and random affiliation, when holding degree distributions of boards and directors fixed, and simulating from the largest connected component, which corresponds to the results in the right column of Panel C in Table 2. P-values of zero obtain whenever observed distance or clustering are outside the range of simulated values from the counterfactual.

Table 6: Count of $(2,2)$-bicliques, maximal (b,d)-bicliques and edges lost in the unipartite projections

Country	$(2,2)$ bicliques	Maximal (b,d)-bicliques							Edges lost in projection	
		(2,2)	(2,3)	(3,2)	(3,3)	(2,4)	(4,2)	Higher Order	Boards	Directors
Argentina	127	1	1	0	0	1	2	5	15	102
Australia	406	95	26	16	2	1	3	5	206	352
Austria	43	3	2	1	1	2	0	2	14	38
Bahrain	6	3	1	0	0	0	0	0	4	6
Bangladesh	132	10	2	2	2	2	1	7	25	113
Belgium	42	11	4	2	0	1	0	1	21	39
Brazil	376	15	8	5	0	0	0	10	79	250
Canada	2901	462	128	97	14	50	17	82	1297	2110
Chile	101	8	6	6	3	8	2	0	36	79
China	184	26	5	3	0	0	0	4	57	178
Cyprus	75	2	0	0	0	1	0	4	9	52
Denmark	4	4	0	0	0	0	0	0	4	4
Finland	13	7	2	0	0	0	0	0	10	13
France	331	49	10	4	0	7	1	6	98	304
Germany	126	34	3	3	0	1	0	5	55	116
Greece	31	5	2	2	0	3	0	0	12	28
Hong Kong	1487	99	49	17	17	34	11	61	458	1158
India	1921	258	117	50	16	51	12	48	804	1553
Indonesia	44	8	2	1	0	0	0	2	18	35
Ireland	5	2	1	0	0	0	0	0	3	5
Israel	207	36	10	9	6	6	1	9	86	147
Italy	410	42	27	7	3	7	3	13	127	369
Jamaica	35	4	3	1	0	1	0	1	11	32
Japan	223	70	19	3	2	4	1	4	129	196
Jordan	7	1	2	0	0	0	0	0	4	7
Kenya	9	6	1	0	0	0	0	0	7	9
Kuwait	2	2	0	0	0	0	0	0	2	2
Malaysia	811	75	42	21	7	20	4	29	299	625
Mauritius	99	8	4	3	1	4	2	2	23	77
Mexico	254	19	8	11	2	5	0	8	60	215
Netherlands	24	1	0	1	0	0	0	1	8	23
New Zealand	16	7	2	1	0	0	0	0	11	15
Nigeria	37	6	1	0	1	2	0	1	16	33
Norway	53	20	2	1	0	1	0	2	28	52
Oman	25	10	3	0	0	1	0	0	13	24
Pakistan	681	43	10	4	5	13	4	30	135	417
Philippines	1007	29	27	25	13	35	8	50	214	669
Poland	3	3	0	0	0	0	0	0	3	3
Portugal	17	2	2	1	0	1	0	0	7	16
Qatar	2	2	0	0	0	0	0	0	2	2
Russia	223	22	11	6	3	7	2	6	58	191
Saudi Arabia	6	6	0	0	0	0	0	0	4	6
Singapore	345	50	11	12	1	6	2	11	148	287
South Africa	207	35	13	8	0	4	0	3	85	194
Spain	71	10	2	2	0	3	0	3	25	68
Sri Lanka	1231	22	28	16	13	21	8	51	265	595
Sweden	114	38	12	9	0	2	0	1	74	97
Switzerland	14	11	1	0	0	0	0	0	12	14
Taiwan	33	7	5	1	0	1	1	0	16	28
Thailand	522	63	32	14	2	15	4	14	180	451
Turkey	140	10	8	2	2	4	2	6	26	83
United Kingdom	41	29	2	2	0	0	0	0	37	39
United States	2051	466	83	40	5	28	7	40	919	1840

This table shows the count of $(2,2)$-bicliques, maximal (b,d)-bicliques found in the observed network, as well as the number of edges lost in the unipartite projections. Bicliques are groups of maximally connected nodes, which means that they have all possible connections between all members. Maximal (b,d)-bicliques are (b,d)-bicliques that are not subsets of any larger biclique. We count the maximal number of (b,d)-bicliques with $(b,d \geq 2, b+d \leq 6)$, where b and d are the numbers of boards and directors involved in each clique. Higher order bicliques is the sum of the count of (b,d)-bicliques with $(b,d \geq 2, b+d > 6)$.

Table 7: Summary of parametric tests for all simulations

	Panel A: Holding density fixed											
	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	39	11	3	35	15	3						
Degree	0	6	47	1	9	43	8	17	28	9	16	28
Clustering	0	3	50	2	17	34	0	3	50	0	3	50
Distance	30	19	4	33	16	4	4	16	33	5	16	32
Small world		19	34		17	36		18	35		18	35
Random affiliation		3	50		11	42		3	50		3	50
$(2, 2)$-bicliques	0	0	53				0	1	52			
	Panel B: Holding degree distribution of boards fixed											
	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	39	14	0	39	14	0						
Degree	0	7	46	3	17	33	8	17	28	51	2	0
Clustering	0	3	50	10	24	19	0	4	49	0	5	48
Distance	26	22	5	26	20	7	4	12	37	2	12	39
Small world		24	29		20	33		14	39		14	39
Random affiliation		3	50		15	38		5	48		2	51
$(2, 2)$-bicliques	0	1	52				0	2	51			
	Panel C: Holding degree distribution of boards and directors fixed											
	From the whole network						From the giant component					
	Boards			Directors			Boards			Directors		
	<	=	>	<	=	>	<	=	>	<	=	>
Giant Component	43	10	0	44	9	0						
Degree	16	25	12	3	19	31	50	3	0	51	2	0
Clustering	5	22	26	41	9	3	1	17	35	5	10	38
Distance	7	23	23	7	25	21	0	7	46	0	6	47
Small world		21	32		5	48		9	44		8	45
Random affiliation		16	37		7	46		8	45		6	47
$(2, 2)$-bicliques	0	2	51				0	3	50			

This table shows summary statistics for parametric tests performed when the density of the data is held fixed in the simulations (Panel A), when imposing on the simulations the degrees distribution of the boards (Panel B), and when both the degree distribution of the boards and directors is imposed on the counterfactual (Panel C). The table shows the number of countries (out of a total of 53 countries with large enough networks) that have simulated measures that are smaller ("<"), not significantly different ("=") or greater (">") than the actual observed measures. Equality corresponds to a bilateral test at 10%, while inequalities correspond to unilateral tests at 5%. Simulations are performed from the entire network in the left half of the table and with data from the giant component only in the right half of the table. Largest component compares the size of actual with simulated giant components. Degree, distance and clustering are as shown in Appendix A. Small world refers to the Kodde & Palm (1986) test of equal distance and excess clustering described in Appendix D.1, while random affiliation refers to a joint Wald test for distance and clustering described in Appendix C.1. Tests for individual measures are based on z-scores and the assumption of normality.

Table 8: Redundant links and randomness of network

	Boards		Directors	
	Distance	Clustering	Distance	Clustering
Intercept	1.12***	−0.56	0.52	1.47*
	(0.33)	(0.47)	(0.40)	(0.85)
Redundant Links	1.19***	1.02***	0.84***	0.95***
	(0.10)	(0.15)	(0.06)	(0.13)
Size	−0.60***	−0.38**	−0.25***	−0.54***
	(0.12)	(0.18)	(0.08)	(0.17)
R^2	0.84	0.70	0.88	0.61
Adj. R^2	0.83	0.69	0.87	0.59
Num. obs.	53	53	53	53

This table shows regression results of the log of the absolute value of the z-scores of distance and clustering for board and director projections on the (log) number of redundant links in the corresponding projection and on size, measured as the log of the number of corresponding nodes. $^{***}p < 0.01$, $^{*}*p < 0.05$, $^{*}p < 0.1$.

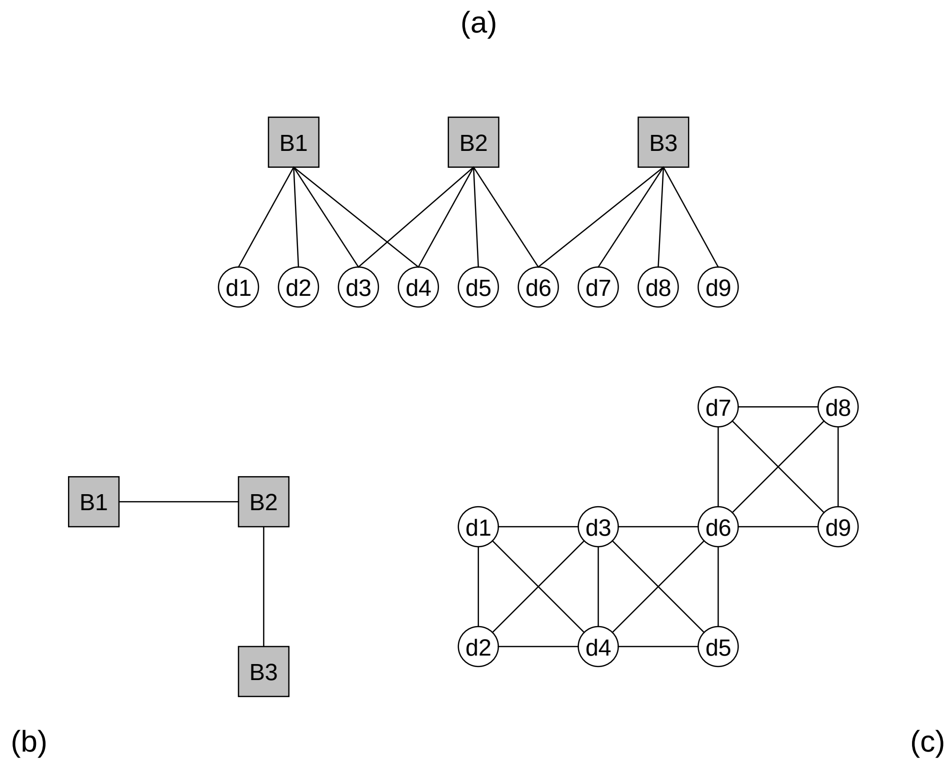


Figure 1: A bipartite network and its two unipartite projections

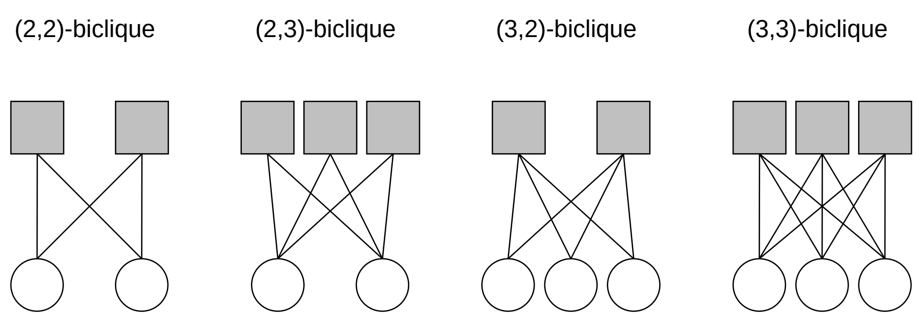
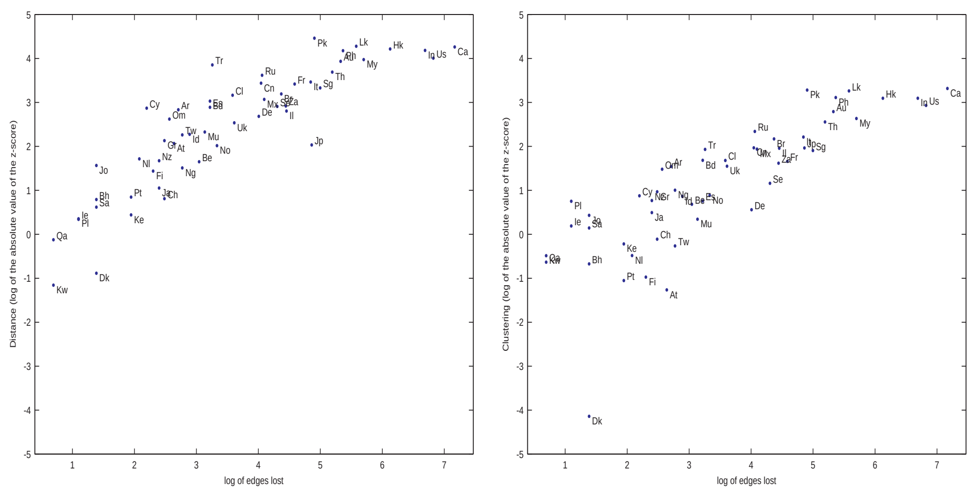


Figure 2: Four instances of (b, d)-bicliques, with $b, d \in \{2, 3\}$.

Figure 3: Distance and clustering vs. number of edges lost in the board projection

(a) Distance (b) Clustering



This figure shows scatter plots of the log of the number of edges lost in the board projection vs. the log of the absolute value of the z-scores of distance in Panel (a) and clustering in Panel (b). z-scores are the difference between observed and average simulated measures, normalized by the standard deviation of the simulations. Both z-scores are based on the counterfactual that imposes the degree distribution of boards and directors in a simulation from the giant component only. Each point corresponds to a country, and is labeled with its two-letter country code.

Appendices

A Formal definitions of network statistics

In this section we detail the computation of the unipartite network measures considered in this paper. Recall that from any bipartite board membership network, the (unipartite) board network obtains by projection on the set of boards B with the convention that two boards are connected when they have at least one common director. Symmetrically, the (unipartite) director network obtains by focusing on the set of directors D and adopting the convention that two directors are connected when they jointly sit on at least one board. A set $S = \{1, \ldots, N\}$ of nodes together with a set g of unordered pairs of nodes (edges) in S thus defines a unipartite network (S, g) , where S is either D or B, and g has been constructed according to the convention just stated.

The *neighborhood* of i consists of the set of nodes to whom i is directly connected, and is denoted $\mathcal{N}_i = \{j \neq i : ij \in g\}$. The size of the neighborhood of i is also called its *degree*, $N_i = |\mathcal{N}_i|$, where $|A|$ denotes the cardinality of set A. The average degree in (S, g) is

$$d = \frac{1}{N} \sum_{i \in S} N_i.$$

For given N, a higher value of average degree will correspond to a denser network. *Density* is defined as the ratio of the degree sum, $\sum_{i \in S} N_i = Nd$, to its largest possible value, $N(N-1)$, which simply amounts to $d/(N-1)$.

Neighbourhoods can be used to examine the extent to which a node's neighbors are neighbors of each other. Neighborhood *clustering* equals the number of edges among i's neighbors divided by the largest possible number of edges among them, according to

$$C_i = \frac{|jk \in g : j, k \in \mathcal{N}_i|}{N_i(N_i - 1)/2}.$$

Clustering is a normalized measure of neighborhood overlap among the neighbors of a given node. The *clustering coefficient* in a network is the average of that statistic taken over all nodes, written as

$$C = \frac{1}{N} \sum_{i \in S} C_i.$$

Beyond direct neighborhoods, one can examine the structure of all indirect relationships existing in a network. A *path* in g is a sequence of edges $i_1 i_2, \cdots, i_{p-1} i_p$ between unordered pairs, without cycles, such that $i_k i_{k+1} \in g, \forall k = 1, \ldots, p-1$. The number of edges in the shortest path between i and j, denoted $d_{i,j}$, is the (geodesic) *distance* between i and j in the network (S, g). It is straightforward to see that $d_{i,j} = 1, \forall j \in \mathcal{N}_i$. By convention $d_{i,j} = \infty$ whenever there is no path connecting i and j. The *largest component* in g is the largest subset $V \subseteq S$ such that $d_{i,j} < \infty$ for all $i, j \in V$ and $d_{i,j} = \infty$ for all $i \in V, j \notin V$. *Average distance* is the average taken over all reachable pairs. If the focus is on the largest component only, then the average is taken for nodes in V only, so that

$$D = \frac{1}{|V|(|V| - 1)} \sum_{i,j \in V, i \neq j} d_{i,j}.$$

B Random network algorithms

For each one of our three counterfactuals, for either the entire data network or the giant component in the data, the random network simulation algorithm consists of the following steps:

B.1 Generate degree sequences;

B.2 Match directors and boards given the degree sequences;

B.3 Connect the graph (only when focusing on the giant component);

B.4 Swap edges to eliminate any possible departure from randomness created by the earlier stages.

B.1 Generate degree sequences

This first step of the algorithm takes as inputs the degree sequences of boards and directors in the actual data (either the entire network or the giant component) and generates degree sequences that take into account the constraints imposed by each counterfactual.

1. **Holding density fixed**: Keeping the number of boards, directors and edges identical to their values in the data we allocate edges randomly to the boards and directors, making sure each board and director has a degree of at least 1 (every director sits on at least one board, and every board has at least one seat).

2. **Holding degree distribution of boards fixed**: Keeping the number of boards, directors and edges identical to their values in the data, and keeping the degree sequence of the boards in the data, we allocate edges randomly to the directors, making sure each director has a degree of at least 1.

3. **Holding degree distribution of boards and directors fixed**: In this case there is no simulation involved, since the degree sequences of the boards and directors are the ones observed in the data.

B.2 Match directors

Given the simulated degree sequences of boards and directors from the previous step, we create a list of boards and a list of directors. The number of copies of each board (director) on the board (director) list equals the simulated degree of this board (director), so the two lists have the same length, equal to the degree sum (twice the number of edges). After a permutation of the list of directors, we match boards and directors: the first director on the (shuffled) director list with the first board on the board list, and so on. This can produce multiple occurrences of the same edge, i.e., a director sitting twice on the same board, which we eliminate by random edge swapping. Suppose there are two copies of B_1d_1 in the simulated ege list. Pick a random edge, say B_2d_2, and one of the copies of B_1d_1, and replace them by the new edges B_1d_2 and B_2d_1. This swap preserves the degrees of the 2 boards and 2 directors involved and has done away with the duplicate. We continue swapping edges until all duplicates are removed from the edge list.

B.3 Connect the graph

At the end of the previous stage we have a random bipartite network but not necessarily a connected one. Whenever we simulate from the giant component, we want a connected network. Again we resort to edge swapping. However, this time we swap edges across different connected components of the network. Suppose B_1d_1 and B_2d_2 are two random edges belonging to two different connected components in the simulated bipartite network. As in the previous stage, we remove B_1d_1 and B_2d_2 from the edge list and replace them with B_1d_2 and B_2d_1. This creates two links between the formerly disconnected components, while preserving the degrees. We iterate this procedure until the network is connected.

B.4 Swap edges

Finally, to get rid of any structure we might have introduced during the previous stages, we conduct a large number of rounds of edge swapping, considering each edge twice on average. Each time we pick two random edges, swap their extremities and we accept this swap provided that it does not create multiple occurrences of edges, and does not disconnect the graph when we simulate from the giant component.

C Random affiliation test

In this appendix, we provide further details on the random affiliation test, which corresponds to the hypothesis that the actual values of distance and clustering are both compatible with the distribution of distance and clustering observed for our counterfactual network. In other words, we want to check that the observed network is not unusual for the type of network we consider.

C.1 Test under the assumption of joint normality

Under the null of random affiliation, and under the assumption of joint normality of distance D and clustering C of the counterfactual networks, the joint test for this hypothesis is the Wald test, $W = (\bar{\gamma} - \mu)'\Sigma^{-1}(\bar{\gamma} - \mu)$, which is distributed $W \sim \chi^2_{[2]}$ where $\bar{\gamma} = (\bar{d}, \bar{c})$ are the actual values of distance and clustering observed in the data, $\mu = (\mu_d, \mu_c)$, where μ_d and μ_c are, respectively, the mean of distance and clustering, and $\Sigma = Cov(D, C)$.

The non-rejection region of the test is

$$R = \left\{ \gamma = (d, c) : (\gamma - \mu)'\Sigma^{-1}(\gamma - \mu) \leq \chi^2_{[2],1-\alpha} \right\},$$

where $\chi^2_{[2],1-\alpha}$ is the critical value at significance level α, which corresponds to an elliptically shaped contour plot of the bivariate normal distribution. Another way of thinking about this is to express the non-rejection region as a function of the density of distance and clustering:

$$R = \{ \gamma = (d, c) : \phi_2(\gamma, \mu, \Sigma) > \phi_\alpha \},$$

where $\phi_2(\gamma, \mu, \Sigma)$ is the Gaussian density with mean μ and variance covariance Σ, and $\phi_\alpha = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}\chi^2_{[2],1-\alpha} \right\}$, which corresponds to the α quantile of ϕ_2. Note that, for the bivariate Gaussian, R corresponds to the highest density region

(HDR), defined in Hyndman (1996) as the smallest possible region that covers a specified probability. These HDRs provide an attractive alternative to the traditional symmetric confidence interval around the mean, especially in multivariate possibly non-Gaussian set-ups. Figure E.1, Panel (a) shows the non-rejection area when distance and clustering have a correlation of 0.5. This is the traditional ellipse, which corresponds to the level curve of the bivariate normal density.

C.2 Nonparametric test

When one is not willing to assume joint normality, it is nonetheless possible to build a test that relies on the comparison of observed values of distance and clustering with the level curves of the joint density of distance and clustering under the null of random affiliation. Thus, the elliptical level curve of the Gaussian is replaced by the level curves derived from the density. The non-rejection region of the test depends on $f(\gamma)$, the joint density of C and D, and f_α, its α quantile:

$$R = \{\gamma = (d, c) : f(\gamma) > f_\alpha\}.$$

In order to compute level curves of the joint distribution of distance and clustering, $f_{D,C}(d, c)$, which is unknown, we estimate it nonparametrically using a kernel:

$$\hat{f}(d, c) = \sum_i K_{h_d}\left((d - d_i)/h_d\right) K_{h_c}\left((c - c_i)/h_c\right),$$

where $K_h(x)$ is a Gaussian kernel with bandwidth h, h_c, and h_d are optimal bandwidths chosen by cross-validation. The non-rejection region of the test is

$$R\left(\hat{f}(d, c)\right) = \left\{(d, c) : \hat{f}(d, c) \geq \hat{f}_\alpha(d, c)\right\},$$

where $\hat{f}_\alpha(d, c)$ is the largest constant such that $P\left((D, C) \in R(\hat{f}(d, c))\right) \geq 1 - \alpha$. We can reject the null hypothesis at 5%, whenever observed values of distance and clustering $(\bar{d}, \bar{c})$ fall outside the 95% HDR, i.e. whenever $\hat{f}(\bar{d}, \bar{c}) < \hat{f}_\alpha(d, c)$. We implement the kernel estimated using R package *np* of Hayfield & Racine (2008), and we use the leave-one-out kernel to estimate the quantile of the density under the null, and the usual kernel estimator when we evaluate the density at the observed data points.

D Small world test

In this appendix, we provide further details on the small worlds test, which corresponds to the hypothesis that the actual values of distance is compatible with the distribution of distance, while the actual value of clustering is either compatible with the counterfactual network or greater. In other words, we want to check that distance in the observed network is not "unusual" for the type of network we consider, while clustering is not "unusually" lower than in the counterfactual.

D.1 Test under the assumption of joint normality

Under the assumption of normality, we rely on the test of Kodde & Palm (1986). The idea of the test is to evaluate the distance in the metric of $\Sigma = Cov(C, D)$ between $\bar{\gamma} = (\bar{d}, \bar{c})$, the observed values of distance and clustering and $\tilde{\gamma} = (\tilde{d}, \tilde{c})$, the projection of the observed values of distance and clustering onto the feasible set under H_0. Distance is projected orthogonally on H_0 as $\tilde{d} = \mu_d$, while for clustering the projection is:

$$\tilde{c} = \min_{x \geq 0} \left(\frac{\bar{c} - x - \mu_{c|d}}{\sigma_{c|d}} \right)^2 ,$$

which yields the closed-form solution $\tilde{c} = \max(\bar{c} - \mu_{c|d}, 0)$, where $\mu_{c|d} = \mu_c - \sigma_{cd}/\sigma_d^2(\bar{d} - \mu_d)$, and $\sigma_{c|d}^2 = \sigma_c^2 - \sigma_{cd}/\sigma_d^2$ are, respectively, the conditional mean and variance of a Gaussian, with σ_d^2, σ_c^2 and σ_{cd}, respectively the variance of distance, clustering and the covariance between them. Thus we have two possible cases for

$\tilde{\gamma} = (\mu_d, \tilde{d})$. The statistic is given by

$$KP = (\bar{\gamma} - \tilde{\gamma})' \, \Sigma^{-1} \, (\bar{\gamma} - \tilde{\gamma}) = \underbrace{\left(\frac{\bar{d} - \mu_d}{\sigma_d} \right)^2}_{\chi^2_{[1]}} + \underbrace{\left(\frac{\bar{c} - \tilde{c} - \mu_{c|d}}{\sigma_{c|d}} \right)^2}_{\text{mixture of } \chi^2_{[1]} \text{ and } 0},$$

where the quadratic form has been decomposed using the partitioned inverse of Σ, into the marginal of D, and the conditional of C, given D. Thus the Wald statistic is decomposed into a univariate statistic for distance, which follows a $\chi^2_{[1]}$ distribution, and a univariate statistic for clustering, conditional on distance, which follows a mixture of point mass at zero (when $\tilde{c} = \bar{c} - \mu_{c|d}$) and a $\chi^2_{[1]}$ distribution (when $\tilde{c} = 0$). Essentially the idea of the test is that if, conditional on distance, clustering is above its mean, then in the test statistic, clustering is replaced by its conditional mean, otherwise, it is set to zero.

Two cases can be distinguished. First, if observed clustering is above its conditional mean, then the second term cancels out, and the statistic boils down to a univariate statistic based on distance: $KP = \left(\frac{\bar{d} - \mu_d}{\sigma_d} \right)^2$. This corresponds to the idea that if clustering is in the non-rejection area, the statistic relies exclusively on distance. However, if clustering is below its conditional mean, the second term is equal to a univariate statistic of the difference between clustering and its conditional mean, which follows a $\chi^2_{[1]}$ distribution and is independent of the distribution of the first term. In that case, the two independent $\chi^2_{[1]}$ terms combine into the traditional Wald test for two equalities, $KP = W$, which follows a $\chi^2_{[2]}$ distribution:

$$KP = \begin{cases} \left(\frac{\bar{d} - \mu_d}{\sigma_d} \right)^2 & \text{if } \bar{c} > \mu_{c|d}, \\ \left(\frac{\bar{d} - \mu_d}{\sigma_d} \right)^2 + \left(\frac{\bar{c} - \mu_{c|d}}{\sigma_{c|d}} \right)^2 & \text{if } \bar{c} < \mu_{c|d}. \end{cases}$$

The mixing probability is given by the conditional probability that clustering exceeds its conditional mean, $P\left(\bar{c} > \mu_{c|d} | C = \bar{c}\right) = \Phi(0) = \frac{1}{2}$, where $\Phi(x)$ is the cumulative density function of the normal, and thus $KP \sim \frac{1}{2}\chi^2_{[1]} + \frac{1}{2}\chi^2_{[2]}$. The non-rejection region of the test is

$$R = \left\{ \gamma = (d,c) : (\gamma - \tilde{\gamma})' \, \Sigma^{-1} \, (\gamma - \tilde{\gamma}) \leq \chi^2_{[1-2],1-\alpha} \right\},$$

where $\chi^2_{[1-2],1-\alpha}$ is the critical value at significance level α of the mixture of χ^2 distributions. Another way of thinking about this is to express the non-rejection region as a function of the density of distance and clustering:

$$R = \{ \gamma = (d,c) : \phi_2(\gamma - \tilde{\gamma}, \Sigma) > \phi_\alpha \},$$

where now $\phi_\alpha = \frac{1}{2\pi\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}\chi^2_{[1-2],1-\alpha} \right\}$, and where

$$\begin{aligned}
\phi_2(\gamma - \tilde{\gamma}, \Sigma) &= \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}\left(\frac{d-\mu_d}{\sigma_d}\right)^2 - \frac{1}{2}\left(\frac{c-\tilde{c}-\mu_{c|d}}{\sigma_{c|d}}\right)^2 \right\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left\{ -\frac{1}{2}\left(\frac{d-\mu_d}{\sigma_d}\right)^2 \right\} \frac{1}{\sqrt{2\pi}\sigma_{c|d}} \exp\left\{ -\frac{1}{2}\left(\frac{c-\tilde{c}-\mu_{c|d}}{\sigma_{c|d}}\right)^2 \right\} \\
&= \phi_2([d, \mu_{c|d}], \mu, \Sigma)\mathbb{1}_{\{c>\mu_{c|d}\}} + \phi_2([d,c],\mu,\Sigma)\mathbb{1}_{\{c<\mu_{c|d}\}}
\end{aligned}$$

Thus we can write

$$R = \left\{ (d,c) : \phi_2([d, \mu_{c|d}], \mu, \Sigma)\mathbb{1}_{\{c>\mu_{c|d}\}} + \phi_2([d,c],\mu,\Sigma)\mathbb{1}_{\{c\leq\mu_{c|d}\}} > \phi_\alpha \right\}.$$

Figure E.1, Panel (b) shows the 5% non-rejection area when distance and clustering have a correlation of 0.5. The non-rejection area is a combination of two areas, depending on whether clustering is above or below its conditional mean. When clustering is above its conditional mean, the non-rejection area is between two vertical lines, since in that case, the null hypothesis can only be rejected when distance is significantly different from simulated distance, i.e. to the right or the left of the vertical lines. When clustering is below its conditional mean, the area has a similar shape as in the random affiliation case, which corresponds to the contour of the bivariate normal density.

D.2 Nonparametric test

We now combine the ideas of Kodde & Palm (1986) with the technique of Hyndman (1996) to derive a nonparametric version of the small world test. The idea is that, as in Kodde & Palm (1986), there are two possible cases, depending on whether observed clustering is above or below its conditional mean. While in the Gaussian case, there is a closed-form expression for the mean of clustering C conditional on distance D, we now rely instead on a nonparametric estimate of the conditional mean, as follows:

$$\hat{E}[C|D] = \frac{\sum_i w_i c_i}{\sum_i w_i},$$

where $w_i = K_{\hat{h}_d}\left((d - d_i)/\hat{h}_d\right)$ and $\hat{h}_d$ is the optimal bandwidth chosen by cross-validation. Depending on whether observed clustering lies above or below its estimated conditional mean, we use either the joint distribution of distance and clustering estimated as above by $\hat{f}(d, c)$, or the same joint distribution, but where clustering is evaluated at conditional mean, $\hat{f}(d, \hat{E}[C|D = d])$. Thus the non-rejection region is

$$R = \{(d,c) : \mathbb{1}_{\{c > \hat{E}[C|D=\bar{d}]\}} \hat{f}\left(d, \hat{E}[C|D = \bar{d}]\right) + \mathbb{1}_{\{c \leq \hat{E}[C|D=\bar{d}]\}} \hat{f}\,(d,c) > \hat{f}_\alpha(d,c)\}$$

where $\hat{f}_\alpha(d,c)$ is the largest constant such that $P\left((D,C) \in R(\hat{f}(d,c))\right) \geq 1 - \alpha$.

E Monte Carlo simulation for nonparametric tests

In this Appendix we briefly present the results of a Monte Carlo study that we carry out to check that the size and power of the tests presented in Appendices C and D are satisfactory. The results are shown in Table 9. We analyze the parametric Wald and Kodde & Palm (1986) tests, as well as their nonparametric versions, which correspond respectively to the random affiliation and small world hypotheses. We consider two alternative distributional assumptions. In the case of normal data (Panel A), the parametric tests are exact, since they are derived under the assumption of normality. In the case with data that follows a Student t distribution with 6 degrees of freedom (Panel B) with significantly fatter tails than the normal, the parametric tests are no longer exact. We consider three different scenarios in terms of the correlation of the data: $\rho = -0.5, 0, 0.5$.

We first check the size of the test, i.e. we make sure that a test with nominal size (significance level) of 5% rejects the null hypothesis 5% of the time, and then we examine the power of the test, which is its ability to reject the null hypothesis when it is wrong. In the case of normally distributed data, the size is very close to the nominal size of 5%, and the difference between the nonparametric test and the Wald and Kodde and Palm tests are very small. With Student t data, the size of the nonparametric tests is still very close to the nominal size, while, as expected, the size of the Wald χ^2 test and the Kodde is off by 0.8% to 1.5%, and consistently above nominal size, which means that the test significantly over-rejects. This shows that the nonparametric tests perform nearly as well as the parametric ones with Gaussian data, and they keep working well with non-Gaussian data, when instead the performance of the standard tests deteriorates.

We also investigate the power of the tests, and as alternative hypotheses we consider that the means are 1, 2 or 3 standard deviations below the mean for both distance and clustering. The differences between the nonparametric tests and their classical counterparts are usually in the second decimal when the data is Gaussian. With Student t data, the power of the nonparametric tests is slightly lower than in the Gaussian case, but it remains high. Given that the parametric tests systematically over-reject, it does not make sense to look at power in that case.

Thus the nonparametric tests perform as well as the parametric ones when the data is normal, and unlike parametric tests, they continue to perform well under deviations from normality.

Table 9: Monte Carlo simulations of parametric and nonparametric random affiliation and small world tests

Panel A: Gaussian					
		Random Affiliation		Small World	
		Non Param	Param	Non Param	Param
Size					
$\rho = -0.5$		5.03	5.01	5.01	5.01
$\rho = 0$		5.02	5.06	4.95	5.03
$\rho = 0.5$		5.01	5.01	4.88	5.01
Distance	Clustering	Power ($\rho = 0$)			
μ_d	$\mu_c - \sigma_c$	13.19	13.28	17.51	17.77
	$\mu_c - 2\sigma_c$	41.20	41.62	48.55	49.19
	$\mu_c - 3\sigma_c$	76.56	77.09	81.90	82.42
$\mu_d - \sigma_d$	μ_c	13.33	13.38	14.06	14.33
	$\mu_c - \sigma_c$	22.53	22.71	28.01	28.53
	$\mu_c - 2\sigma_c$	50.02	50.44	57.21	57.94
	$\mu_c - 3\sigma_c$	81.03	81.60	85.64	86.22
$\mu_d - 2\sigma_d$	μ_c	41.26	41.63	43.71	44.41
	$\mu_c - \sigma_c$	50.12	50.50	56.56	57.24
	$\mu_c - 2\sigma_c$	71.35	71.83	77.11	77.78
	$\mu_c - 3\sigma_c$	90.55	90.89	93.25	93.58
$\mu_d - 3\sigma_d$	μ_c	76.69	77.16	79.12	79.68
	$\mu_c - \sigma_c$	81.24	81.62	85.30	85.72
	$\mu_c - 2\sigma_c$	90.56	90.83	93.28	93.54
	$\mu_c - 3\sigma_c$	97.35	97.47	98.27	98.39

Panel B: Student t (df=6)					
		Random Affiliation		Small World	
		Non Param	Param	Non Param	Param
Size					
$\rho = -0.5$		5.05	6.44	4.93	5.83
$\rho = 0$		5.00	6.41	4.93	5.79
$\rho = 0.5$		5.08	6.44	4.97	5.86
Distance	Clustering	Power ($\rho = 0$)			
μ_d	$\mu_c - \sigma_c$	9.25		13.04	
	$\mu_c - 2\sigma_c$	30.96		41.61	
	$\mu_c - 3\sigma_c$	72.34		81.19	
$\mu_d - \sigma_d$	μ_c	9.20		10.43	
	$\mu_c - \sigma_c$	15.09		21.12	
	$\mu_c - 2\sigma_c$	40.11		51.82	
	$\mu_c - 3\sigma_c$	78.09		85.55	
$\mu_d - 2\sigma_d$	μ_c	30.83		37.25	
	$\mu_c - \sigma_c$	40.06		51.06	
	$\mu_c - 2\sigma_c$	65.76		75.88	
	$\mu_c - 3\sigma_c$	89.17		93.30	
$\mu_d - 3\sigma_d$	μ_c	72.23		78.44	
	$\mu_c - \sigma_c$	77.88		84.93	
	$\mu_c - 2\sigma_c$	89.12		93.14	
	$\mu_c - 3\sigma_c$	96.57		97.87	

This table shows the size and power of the parametric (Wald and Kodde-Palm) and nonparametric tests of random affiliation and small worlds. Panel A shows simulations done with jointly normal distance and clustering, while Panel B presents results obtained from Student t data with 6 degrees of freedom. There are three scenarios in terms of the correlation of the data: $\rho = -0.5, 0, 0.5$, and power is analyzed when under the alternative hypothesis the means are 1, 2 or 3 standard deviations below the mean for both clustering and distance.

(a) Random affiliation (b) Small world

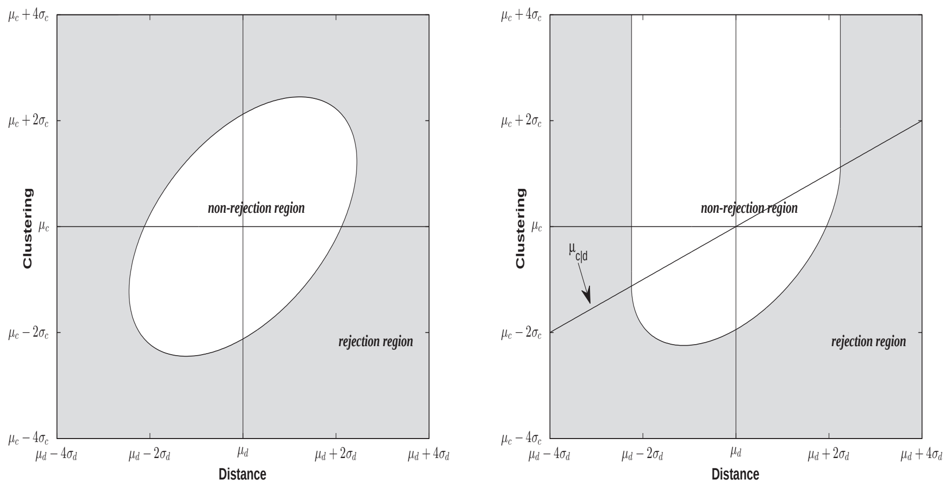


Figure E.1: Non-rejection regions for random affiliation and small world test

This figure shows the 5% non-rejection areas for the random affiliation and for small world tests based on joint normality of distance and clustering. Panel (a) shows the non-rejection region for random affiliation:

$$R = \{\gamma = (d, c) : \phi_2(\gamma, \mu, \Sigma) > \phi_\alpha\},$$

where $\phi_\alpha = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\chi^2_{[2],1-\alpha}\right\}$, and $\chi^2_{[1-2],1-\alpha}$ is the critical value at significance level α of the mixture of χ^2 distributions, and ϕ_2 is the density of the bivariate normal. Panel (b) shows the 5% non-rejection region for the small world test:

$$R = \left\{(d, c) : \phi_2([d, \mu_{c|d}], \mu, \Sigma)\mathbb{1}_{\{c>\mu_{c|d}\}} + \phi_2([d, c], \mu, \Sigma)\mathbb{1}_{\{c\leq\mu_{c|d}\}} > \phi_\alpha\right\},$$

where $\phi_\alpha = \frac{1}{2\pi\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\chi^2_{[1-2],1-\alpha}\right\}$. The line with slope 0.5 corresponds to the mean of clustering conditional on distance. The plots are shown in a range of up to 4 standard deviations, σ_d, and σ_c below and above the means, μ_d and μ_c of distance and clustering.