

*The Small World Network Structure of Boards of
Directors*

Conyon, M. J. and Muldoon, M. R.

2005

MIMS EPrint: **2005.15**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

The Small World Network Structure of Boards of Directors

Martin J. Conyon^{1*} and Mark R. Muldoon²

¹The Wharton School, University of Pennsylvania, USA

²Department of Mathematics, University of Manchester, UK

May 2004

Abstract

In this paper we present a random graph model to explain the network structure of boards of directors. We investigate the conditions under which corporate boards can be said to be a “small-world”. Our empirical results show that the random graph model is remarkably good at explaining board structure and connectedness in the United States, the United Kingdom and Germany. Although there are small-world traits such as “clustering” and “short-paths” in the corporate world, they are no more pronounced than would be expected by chance in a statistically similar, but randomly assembled corporate universe. Finally, our results show the existence of positive degree correlation: directors who sit on many boards do so in the company of other directors who sit on many boards. This result helps explain the distribution of board interlocks.

1 Introduction

The board of directors is central to corporate governance – it is the prime decision making body in the public corporation ((Hermalin & Weisbach (2003)). An important feature of such boards is that they are often connected to each other by means of a shared director. As Fama & Jensen (1983) remark: “Most outside directors of open corporations are either managers of other corporations or important decision agents in other organizations”. Such network connectivity has important economic consequences. For instance, boards that are “interlocked” in this fashion facilitate the diffusion of executive compensation practices – Hallock (1997). In addition, such networks promote the adoption of anti-takeover

***Acknowledgments:** E-mail conyon@wharton.upenn.edu or M.Muldoon@UMIST.ac.uk with comments. We would like to thank Lori Rosenkopf and seminar participants at Manchester University for comments and discussions during the preparation of this paper. We are grateful to Lerong He for research assistance and to the Reginald Jones Center of the Wharton School for financial support.

mechanisms such as poison pills as well as the spread of golden parachutes – Davis & Greve (1997).

This paper contributes to the corporate governance literature in the following ways. First, we present a random graph model of the board of directors. The model illustrates the conditions under which one can identify a “small-world”. The idea of a small-world is a potentially ambiguous so to fix ideas we follow the conventions introduced by Newman, Strogatz & Watts (2001). A small-world is characterized by two properties. First, there is high network transitivity or clustering. This is the propensity for vertex pairs (e.g. boards) to be connected if they share a mutual neighbor. Second, path lengths or distances between vertices (e.g. boards) are relatively short. This means that any two vertices can be reached in a small number of steps. However, the presence (or not) of a small-world depends critically on the standards to which the computed statistics (e.g. the clustering coefficient) are compared. In short, when one hears that a particular social world is small, one should ask; “Small compared to what?” We address this issue.

Our second contribution is to test the random graph model using board data from three separate economies: the United States, the United Kingdom and Germany. These countries have very different corporate governance arrangements and so provide a rich array to evaluate the predictions arising from the random graph model. Corporate boards in the United States are unitary in structure meaning that insiders (executives) and outsiders (monitors) are members of the same board. Typically insiders represent a tiny fraction of the total board and the posts of chairperson and chief executive officer are usually held by the same person – Core, Holthausen & Larcker (1999). Boards in the United Kingdom are also unitary in structure comprising executive (inside) and non-executive (outside) members. However, the fraction of insiders on the board is generally larger compared with the United States and the leadership roles of CEO and chairperson are typically held by different people – Conyon & Murphy (2000). German companies are governed by a two-tier board structure. The first tier is the supervisory board and is made up of shareholder and employee representatives as well as other stakeholders such as bank representatives. The supervisory board provides the decision control function in German companies. The second tier is the management board which is equivalent to the executive directors in the United Kingdom or the United States. The chairperson of the management board is not a member of the supervisory board – Franks & Mayer (2001).¹ We test whether the random graph model we develop in Section 2 adequately describes the social network of corporate boards in these three diverse economies.

Our paper differs from previous studies in a number of respects. Most previous research on the network structure of boards has focused on the phenomenon of the interlocking director rather than on small-world structures *per se*. Dooley

¹Hermalin & Weisbach (2003) survey the economics of the board of directors. Recent board studies using UK data include Dahya, McConnell & Travlos (2002), Young (2000), and Peasnell, Pope & Young (2003). For Germany see Franks & Mayer (2001) and Conyon & Schwalbach (2000).

(1969) finds that interlocking boards are common place in US firms in 1935 and 1965. Similarly, Hallock (1997) presents evidence, based on a sample of 602 US companies in 1992, that 20 percent are any-employee interlocked (defined as a current or retired employee from firm A sits on the board of firm B and a current or retired employee from firm B who sits on the board of firm A) and 8 percent of firms are CEO interlocked. Hallock (1997) also demonstrates a positive correlation between CEO pay and the presence of an interlocked board. Booth & Deli (1996) report the average number of off-board directorships held by CEOs in over 400 US companies in 1989, is 1.87. The median and maximum values are 2 and 8 respectively. They also show that the number of outside directorships held by the CEO is negatively correlated with firm growth opportunities measured by Tobin's Q (because the opportunity cost of spending time at another firm is high). Studies of this type illustrate the importance of interlocking directors in the US but do not explicitly focus on the small-world structure of the corporate board network.

Davis, Yoo & Baker (2003) compare the small world characteristics (*clustering* and *path length* measures) to those one would expect from a family of random graphs originally introduced in Erdős & Rényi (1959) and Erdős & Rényi (1960). But, as we will illustrate in Section 2 below, the structure of corporate boards imposes constraints on the sorts of social network to which they can give rise. For example, the finiteness of the working year means that even the most energetic of directors can serve on, at most, a few dozen boards. More generally, the distributions of *degree* in the network of corporate directorship—the distributions of the number of directors on a board and the number of boards on which a typical director serves—influence strongly the expected values of small-world measures.²

The rest of this paper is organized as follows: Section 2 outlines a random graph model of the network of corporate boards; Section 3 describes the data and provides empirical results and finally, in Section 4 we offer some concluding remarks.

2 A random graph model of corporate board networks

In this section we define terms and review briefly some important features of graph theory as applied to social networks. We also introduce several families of random graphs designed to share properties with empirically observed social networks. Finally, we introduce the analytic machinery of probability generating

²A small set of papers exist on small-world applications to other corporate governance issues. Kogut & Walker (2001) study the small world of ownership in Germany between 1993 and 1997. The focus of their paper is on ties between firms brought about by a common owner rather than a shared director. Similarly, Baum, Shiplov & Rowley (2003) examine the small-world of Canadian investment banks and their membership of underwriting syndicates rather than shared directors. These studies also share a common feature that they compare small-world characteristics to Erdős & Rényi (1959) type random graphs. As noted we outline an alternative standard.

functions, a circle of techniques that will enable us to decide whether the world of corporate directors appears “smaller” than one would expect.

2.1 Basic terminology

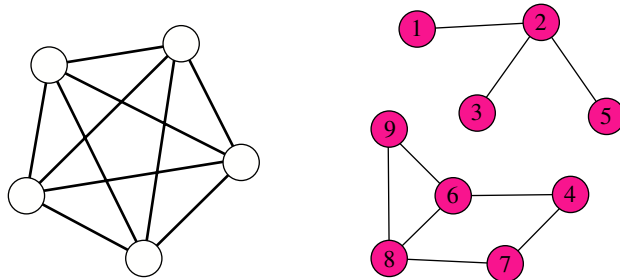


Figure 1: *Two simple graphs. The one on the left is complete while the one on the right has two connected components.*

A *network* (or *graph*) is a set of items termed *vertices* (or *nodes*) with connections between them called *edges*. At this level of abstraction graphs can be used to represent a vast range of phenomena (from networks of marriage alliances in Renaissance Florence to interactions in gene regulatory networks to connections between computers in the internet or between the pages of the World Wide Web). We restrict our attention to networks derived from the world of corporate boards of directors and adopt the following conventions: our nodes will be of two types, either boards or the directors who sit on them. Edges will represent, among other things, membership of a board. We will denote the number of vertices in a graph by N and the number of edges by M . Two vertices will be said to be *adjacent* if they are connected by an edge so, for example, the nodes numbered 1 and 2 in Figure 1 are adjacent, while those numbered 3 and 5 are not.

The concept of *degree* will prove especially important. It is the number of edges connected to a vertex, so all the vertices in the left part of Figure 1 have degree 5. We will use the symbol k_j for the degree of the j -th vertex and z for the average degree of all the vertices in a graph. Next we define the *connected component* associated with a vertex: it is that part of the graph consisting of the vertex itself and all those others that can be reached by paths running along the edges of the graph. In Figure 1 the graph at left has a single connected component while the one at right, the one with colored, numbered vertices, has two connected components: one associated with vertices 1, 2, 3 and 5 and another associated with vertices 4, 6, 7, 8 and 9.

A *geodesic path* (or just *geodesic*) between two vertices is a shortest path (in the sense of traversing the fewest edges) that connects them. Note that there may be more than one geodesic path between a pair of vertices, as in the right

part of Figure 1 where both the paths 4-6-8 and 4-7-8 connect vertices 4 and 8 and both paths involve two edges. The *distance* between two vertices is the number of edges in a geodesic path connecting them. The notion of distance here is intrinsic to the graph and does not depend on any particular plot or image of the graph. The distance between two nodes is undefined if they lie in different connected components.

A graph is said to be *simple* if it has, at most, one edge between each pair of vertices and no edges connecting a vertex to itself. And, finally, a simple graph is called *complete* if it has an edge between every pair of vertices (such as the one at left in Figure 1). A much more extensive discussion of graphs, their representation, manipulation and application to the social sciences appears in Wasserman & Faust (1994).

2.2 Graphs of boards and directors

Data about boards of directors present an immediate problem: how should one draw a graph to represent it? The issue is that one could treat the “board” as the basic unit of analysis and form a graph whose vertices represent boards and whose edges represent shared directors. But alternatively, one could focus on the “board” and make a (generally much larger) graph whose vertices represent directors and whose edges represent shared board memberships. There is no obvious way to choose between these two representations and so most authors simply analyze both. Indeed, we will do the same, reporting results for a “board graph” and a “director graph” according to whether the vertices represent boards or directors.

But the ambiguity about representation arises naturally from the structure of the data: there really are two sorts of social entities here, the directors and the boards, and the network’s edges represent membership of the former in the latter. The most natural representation of such a network, sometimes called an *affiliation network*³, a graph with two sorts of vertices—one each for boards and directors—that has edges connecting directors with the boards on which they sit. The result is an example of a *bipartite* graph: one whose vertices can be divided into two distinct sets and whose edges only make connections between the two sets. See Figure 2 for an example.

The board and director graphs mentioned above now appear as “projections” of the bipartite graph onto one of its two sets of vertices. There is a small potential ambiguity in passing from the bipartite graph to these projections. Suppose that a pair of boards share two or more directors: when preparing the board-vertex projection should one draw multiple edges between these two boards? In practice this issue affects only a tiny fraction⁴ of the edges and so

³Alternative vocabulary abounds: affiliation networks have also been called *membership networks* or *hypernetworks* and are examples of *two-mode* networks. See Chapter 8 of Wasserman & Faust (1994) for an extensive discussion of such networks and their representation.

⁴In the board-vertex projection of the network of UK directorships discussed in Section 3 around 92.4% of the connections between boards involved only a single shared director and 99.5% of interlocked boards shared 3 or fewer directors.

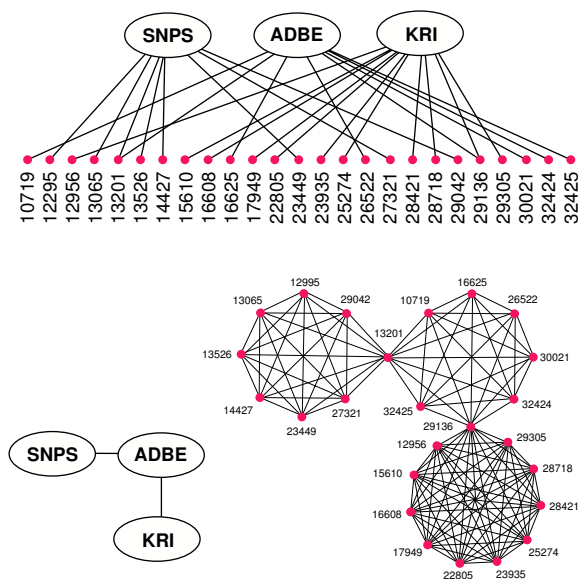


Figure 2: *Bipartite graph representing the neighborhood of Adobe Inc.*

we ignore it, drawing only a single edge for each interlock and so producing projections that are simple graphs.

The bipartite representation of the social network of boards of directors turns out to be important. We will analyze the two projections separately, but when addressing whether there is a small-world or not, the bipartite structure of the underlying affiliation network will bear strongly on the related question “Small compared to what?”

Figure 2 shows a small part of the social network of corporate governance connected to the board of Adobe Systems Inc., a software house whose board has eight members. The top part of the figure is a bipartite graph that includes data for Adobe and those boards with which it is interlocked (that is, those boards with whom Adobe shares a director): the three elliptical nodes at the very top of the figure represent boards and are labelled with ticker symbols. In addition to the one for Adobe (ticker symbol ADBE, a member of the NASDAQ) these are nodes for Synopsis Inc. (SNPS, a member of the NASDAQ) and Knight Ridder Inc. (KRI, a member of the NYSE). The rest of the nodes in the graph are shown as small filled circles and are labelled with numbers. They represent directors and the numbers are IDs assigned by the *The Corporate Library*. To be concrete, director (vertex) 29136 is both a member of ADBE and KRI while director 13201 is a member of both ADBE and SNPS.

The lower part of the figure shows the two projections of the bipartite graph. The one appearing in the lower left part of the figure has the boards as its vertices and edges connecting boards that share a director. The graph at lower

right has directors for vertices and includes an edge between two directors if they sit on the same board. This second projection is striking: it consists of three complete graphs—one for each board—linked together by the two directors (out of a total population of 25) responsible for the interlocks. As we will see below, the fact that the director graph consists of many complete graphs linked together by comparatively rare shared directors will mean that the network of directors appears “highly clustered” in the sense used in the small-world literature. But this high degree of clustering arises automatically because of the way one constructs the board projection: deciding whether the clustering we actually observe is in any sense *unexpectedly* large will be one of our main concerns.

2.3 Characterizing the small-world

In ordinary speech one says “It’s a small world” when, for example, a stranger met in an airport turns out to know one of your grade school teacher’s kids. The surprise in such encounters has two aspects: it’s odd enough to find that a short chain of acquaintance can reach such a long way (you to teacher to kid to far-flung stranger), but seems odder still when one reflects that one’s circle of acquaintance is strongly “clustered” or “transitive”. That is, there is a strong tendency for one’s acquaintances to be acquainted with each other.

Social networks represented as graphs permit one to quantify the two components of the small-world surprise. In this paper we adopt the conventions of Newman (2003) and consider two main kinds of numerical measures: a *characteristic length* L and a measure of *clustering* C . These are defined so that a small-world is characterized by relatively high degree of clustering and by a short characteristic path length. Whether the world of corporate directors really is small depends on the standards used to define “relatively high” and “short” in the previous sentence. As discussed above, the projection onto the directors in Figure 2 is very highly clustered: all the directors belong to medium-sized boards and, within these, are connected to every other director. But this apparently high clustering is just a consequence of the bipartite structure of the corporate social network: below we will introduce analytic machinery that permits one to account for such structural biases.

All of the graphs we study arise as projections of bipartite graphs representing affiliation networks. They may, however, have more than one connected component and usually do. Typically one component is much larger (in the sense of containing a vast majority of the vertices) than all the others. The existence of such a “giant-component” is a much-studied phenomenon in the theory of random graphs (see, for example, Bollobás (2001) or Alon & Spencer (2000)). For the moment we need only observe that the small-world statistics, one of which involves distances between pairs of vertices, only make sense when measured on a connected component. For the rest of this paper we will restrict our attention to the largest connected component and write “graph” when we mean “largest connected component of the graph”.

2.3.1 Typical distances

Half of the small-world phenomenon is that randomly-chosen pairs of vertices turn out to be unexpectedly close to each other. A graph with N vertices contains $N(N-1)/2$ unordered pairs of vertices. If one imagines the vertices to be numbered $1, 2, \dots, N$ and defines $d_{i,j}$ to be the distance between vertex i and j then the simplest way to characterize a typical distance is to take the average

$$L = \frac{2}{N(N-1)} \sum_{i < j} d_{i,j}. \quad (1)$$

The most computationally efficient approaches to calculating distances rely on Dijkstra's Single-Source-Shortest-Path algorithm (see e.g. Cormen, Leiserson & Rivest (2001)), which yields all the distances from a single node to the rest of the vertices. One employs it for each vertex in turn to get a complete list of the $d_{i,j}$. This can be a very computationally expensive business for large graphs, but the networks studied here are sufficiently modest (the projections whose nodes are directors have around 11,000 vertices) that we report exact values for L .

In the network whose nodes are directors L provides a measure of the typical number links in the shortest chain connecting directors: a rumor or a novel idea has to pass through, on average, L intermediaries to get from any one director to any other.

2.3.2 Clustering

There are two commonly used measures of clustering, which is the tendency for a node's neighbors to be neighbors in their own right: both are designed to lie between 0 and 1. A clustering coefficient of 1 means that the graph includes an edge between every pair of vertices (that is, it is a complete graph) while a clustering coefficient of 0 means that the graph does not contain any triples of vertices that are linked into a triangle. The first of the clustering measures, used for example by Watts and Strogatz in Watts (1999), begins by defining a clustering coefficient for each vertex. Use C_j to denote the ratio

$$\begin{aligned} C_j &= \frac{\text{(number of edges between neighbors that actually exist)}}{\text{(number of edges that could exist)}} \\ &= \frac{\text{(number of edges between neighbors of vertex } j)}{k_j(k_j - 1)/2} \end{aligned} \quad (2)$$

where k_j is the degree of the j -th vertex. For example, in the graph at the right of Figure 1 node 4 has degree $k_4 = 2$ and so there could be only a single edge between its neighbors (vertices 6 and 7). In fact this edge isn't present, so $C_4 = 0$. By contrast node 6—which has degree $k_6 = 3$ and so could have $3 \times (3 - 1)/2 = 3$ edges between its neighbors—has only one such edge (connecting vertices 8 and 9) and so has $C_6 = 1/3$.

Once all the C_j terms are calculated it is natural to define a clustering coefficient for the entire graph by averaging:

$$C_{WS} = \frac{1}{N} \sum_{j=1}^N C_j \quad (3)$$

Here the subscript WS is included to distinguish the Watts-Strogatz clustering coefficient from an alternative described below. The larger of the two connected components at the right of Figure 1 has 5 vertices with $C_9 = 1$, $C_6 = C_8 = 1/3$ and $C_4 = C_7 = 0$ so it has overall clustering $C_{WS} = (1/5) \times [1 + (1/3) + (1/3) + 0 + 0] = 1/3$.

The other measure used to characterize clustering is a property of the graph as a whole. It is defined by:

$$C_{\Delta} = \frac{3 \times (\text{number of triangles in the graph})}{(\text{number of connected triples})} \quad (4)$$

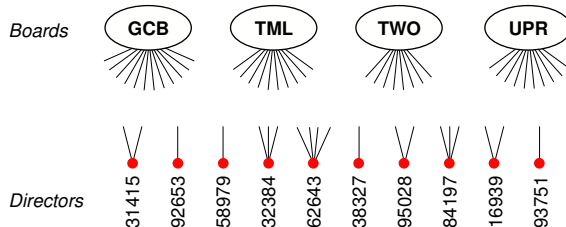
where a *triangle* is a set of three vertices j, k, l in which each vertex is connected to both the other two. A *connected triple* is a set of three vertices j, k, l in which j is connected to k and k is connected to l (though l need not be connected to j). Referring again to the larger of the two connected components at the right of Figure 1, there is only a single triangle (6-8-9), but eight connected triples (6-4-7, 4-6-8, 4-6-9, 8-6-9, 4-7-8, 6-8-9, 7-8-9 and 6-9-8) so the clustering coefficient is $C_{\Delta} = 3 \times 1/8 = 3/8$. As this example demonstrates, the two clustering coefficients are not generally equivalent.

Both clustering coefficients measure the extent to which being a neighbor is a transitive property and, indeed, C_{Δ} is sometimes referred to as the *transitive closure* of the graph. To see why, reconsider the definition above: if the vertices j, k, l form a connected triple then j is connected to k , k is connected to l and, if “is-connected-to” were a strictly transitive relationship one could conclude that j is connected to l . The clustering coefficient C_{Δ} is precisely the proportion of triples for which the transitive conclusion actually holds. The factor of 3 in the numerator is necessary to ensure that $0 \leq C_{\Delta} \leq 1$ as each triangle gives rise to 3 connected triples.

2.3.3 Degree distributions and families of random graphs

The degree of the j -th vertex, k_j , counts, equivalently, either the number of j 's neighbors or the number of edges connected to j . When summarizing a social network one might report the mean or median degree and, as many interesting social networks are far too large to visualize, such summary statistics are often the only means to get a sense of what the graph is like in the large. But they are a bit coarse for the kind of analysis we want to do here: a better tool is the *degree distribution*. As a summary statistic it lists the degrees that occur in the graph and the frequencies with which they appear. Thus the degree distribution for the two-component graph at right in Figure 1 is summarized in Table 1. But one can also interpret the frequencies in the degree distribution as

Stage 1: Generate random degrees for all nodes



Stage 2: Join randomly-selected half-edges

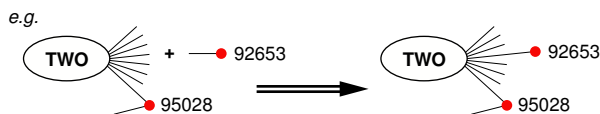


Figure 3: *Generating process for graphs*

probabilities. When viewed in this light an individual graph appears as a single representative of a larger family of graphs, all of whose nodes have degrees drawn from a common distribution. This approach proves extremely fruitful: it will provide the standards to which we compare real social networks when trying to decide whether the world they represent is “small”.

Suppose, for example, that one wanted to construct an ensemble of graphs based on the degree distribution in Table 1. One would begin by fixing the number of nodes N . One would then choose a degree for each node from the degree distribution. At this stage in the construction it is helpful to imagine the j -th node as a ball with k_j spikes sticking out (see Figure 3). The spikes represent half-edges and one assembles a finished graph by joining half-edges at random. The graph construction algorithm sketched here can fail in the sense that one may finish up with a single unmatched half-edge, but this problem is easily repaired. In the first stage of the algorithm, when choosing the degrees for the nodes, one imposes the constraint that the sum of the degrees of all the nodes must be an even number. If that proves not to be the case, one simply starts afresh, choosing more batches of degrees until the constraint is satisfied.

Our real interest is to generate random corporate networks, but this requires only a small extension of the procedure sketched above. One starts with separate degree distributions for the directors and the boards. One then fixes both a total number of boards and a total number of directors and chooses degrees for each. The half-edges now have natural interpretations: those for the boards are unfilled seats while those for the directors are unfulfilled desires for directorships. The problem of avoiding unmatched half-edges is here rather more severe. One needs to ensure that there are exactly as many seats on boards as there

are directors wanting to fill them. But the problem admits a solution similar to the one sketched above: one chooses, say, the degrees for the boards first, then repeatedly generates degrees for the directors until the desired match-up is achieved.

Armed with this algorithm, one could generate many thousands or millions of imaginary corporate worlds that shared a degree distribution with the real object of study. Further, one could measure small-world statistics for each of these imaginary corporate worlds and so obtain typical values (means, say) for the characteristic length L and clustering coefficients C_{WS} and C_{Δ} . And if the real corporate world had a larger C_{Δ} than typically expected, but a smaller L one might reasonably conclude that it was unexpectedly small. This is the approach we take, but we compute expected values for the small-world statistics without generating and analyzing all those random graphs: the necessary theoretical tool, the machinery of probability generating functions, is described next, but before concluding our discussion of random graphs one other family deserves mention, that introduced by the mathematicians Erdős and Rényi.

In their original work they imagined fixing the number of nodes N and then deciding, at random, independently and with fixed probability p , whether each of the graph's $N(N-1)/2$ possible edges exist. In such a graph each node can have as many $(N-1)$ edges—one connecting it to each of the remaining nodes—and in expectation a fraction p of these will exist, so such a graph has mean degree $z = p(N-1)$. More generally, the degree distribution is binomial:

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-k-1}$$

for $k \in \{0, 1, \dots, N-1\}$ and zero otherwise. Here p_k is the probability of finding a node of degree k . Erdős and Rényi were primarily interested in how the qualitative properties of such graphs changed as they held the mean degree z constant, but allowed the number of nodes to tend to infinity. In this limit the degrees of the nodes have a Poisson distribution

$$p_k = \frac{z^k e^{-z}}{k!} \tag{5}$$

As a shorthand we will refer to graphs with Poisson degree distributions as *Poisson random graphs*. The properties of such graphs have been studied extensively, both because of their great mathematical interest (see e.g. Bollobás (2001)) and as models of social networks: early work by Watts and Strogatz, Watts (1999), as well as that of Davis et al. (2003), compare small-world statistics from actual social networks to those expected for random graphs with Poisson degree distribution.

2.3.4 Generating functions

This section draws heavily on the pioneering work of Newman, Strogatz and Watts in Newman et al. (2001) and Newman, Watts & Strogatz (2002) on the

Degree	Frequency
1	3/9
2	3/9
3	3/9

Table 1: Degree distribution for the graph at right in Figure 2.

applications of random graphs with arbitrary degree distributions. See also Wilf (1990) for an expanded discussion of generating functions. The degree distribution associates a probability p_k with each possible value of the degree k , where k is any non-negative integer. Such distributions permit one to construct a *probability generating function*, $G(x)$, which is a function of one variable defined by the infinite sum

$$G(x) = \sum_{k=0}^{\infty} p_k x^k. \quad (6)$$

and it is a general property of generating functions that $G(1) = 1$. For generating functions arising from degree distributions the mean degree is given by

$$z = \langle k \rangle = \sum_{k=0}^{\infty} k p_k = G'(1) \quad (7)$$

Here the angle brackets are the expectation operator. Formulae for higher moments of the distribution, as well for the generating functions for sums of independent samples from distribution are also simply related to $G(x)$. Generating functions arising from empirical degree distributions are, of necessity, finite polynomials as any real data set has a node of highest degree and thus a maximal nonzero p_k . For example, the distribution listed in Table 1 gives rise to the generating function

$$G(x) = \frac{1}{3}x + \frac{1}{3}x^2 + \frac{1}{3}x^3.$$

A bipartite graph of the corporate world has two empirical degree distributions and so gives rise to two generating functions. One, which we shall call $f_0(x)$, generates the degree distribution for the directors. That is, it generates the distribution of the number of boards on which a director sits. For concreteness, say that the frequency with which one finds a director serving on j boards is p_j . Then $f_0(x)$ is given by

$$f_0(x) = \sum_j p_j x^j \quad (8)$$

We shall refer to the other empirical distribution, the one for the degree of boards (that is, the number of directors per board) as $g_0(x)$. So if the empirical frequency for boards made up of k directors is q_k then $g_0(x)$ is

$$g_0(x) = \sum_k q_k x^k \quad (9)$$

Our real objects of interest are the two projections of the bipartite graph: one whose nodes are boards and whose edges represent shared directors (the graph showing board interlocks) and another whose nodes are directors and whose edges connect directors that sit on one (or more) board in common. The charm of the generating function approach is that it permits one to start with the empirical generating functions $f_0(x)$ and $g_0(x)$ and derive expressions for the generating functions for the degree distributions in the projections. We will refer to these derived distributions as the *theoretical degree distributions* to emphasize that they are not measured directly from the data. Instead, they describe the distribution of degrees one would find in random corporate worlds constructed by applying the methods of the previous section to the empirical degree distributions. These theoretical degree distributions are the device by which we obtain expected values of the small-world statistics without having to generate and analyze random graphs.

Suppose now that we are investigating a corporate world in which N directors sit on M boards. Suppose further that the mean number of seats on a board is ν and that the mean number of directorships held is μ . The bipartite graph representing this community has one edge for each seat on a board and so

$$\nu M = (\text{number of seats on boards}) = \mu N.$$

We will refer to the generating function for theoretical degree distribution of the projection onto directors as $G_0(x)$. Newman et al. (2001) show that it is given by

$$G_0(x) = f_0\left(\frac{g'_0(x)}{g'_0(1)}\right) = f_0\left(\frac{1}{\nu}g'_0(x)\right) \quad (10)$$

where $g'_0(1) = \nu$ is the mean board size. As the empirical generating functions $f_0(x)$ and $g_0(x)$ are finite polynomials, $G_0(x)$ is too. By examining its coefficients one can obtain predictions for the frequencies with which vertices of various degrees will appear in the projection: that is, one can obtain the probabilities that define the theoretical degree distribution.

But one can predict other things as well; the expected degree of a vertex in the projection (that is, the expected total number of codirectors for a randomly selected director) is

$$z = G'_0(1) \quad (11)$$

and the expected mean path length in the projection is

$$\langle L \rangle = 1 + \frac{\ln(N/G'_0(1))}{\ln\left(\frac{f''_0(1)}{f'_0(1)}\right)\left(\frac{g''_0(1)}{g'_0(1)}\right)}. \quad (12)$$

Note that the quantity in the denominator involves ratios of the first and second derivatives of the empirical generating functions. Additionally, one can predict the value of one of the two clustering coefficients:

$$\langle C_{\Delta} \rangle = \frac{M}{N} \frac{g'''_0(1)}{G''_0(1)}. \quad (13)$$

The corresponding expressions for the projection whose vertices are boards may be obtained similarly and are

$$\begin{aligned}
F_0(x) &= g_0\left(\frac{f'_0(x)}{f'_0(1)}\right) \\
z &= F'_0(1) \\
\langle L \rangle &= 1 + \frac{\ln(M/F'_0(1))}{\ln\left(\frac{g'_0(1)}{g'_0(1)}\right)\left(\frac{f'_0(1)}{f'_0(1)}\right)} \\
\langle C_\Delta \rangle &= \frac{N}{M} \frac{f''_0(1)}{F''_0(1)}
\end{aligned} \tag{14}$$

In our empirical work below we use equations 10,12, and 13 for the projection whose vertices are directors and equations 14 for the projection whose vertices are boards. We then calculate the actual path lengths and clustering coefficients from the data on the US, Britain and Germany. We compare these to the afore mentioned random graph measures, constructed on the observed degree distribution. Our approach differs from such prior small-world research in the social science literature as that of Kogut & Walker (2001), Davis et al. (2003), Baum et al. (2003) and Schilling & Phelps (2004). These researchers compare actual small-world measures to those arising from an unstructured (i.e. Poisson) random graph. However, the degree distribution of their networks is likely to be measurably different from a Poisson distribution. In consequence, such an approach can lead to the possibility of concluding that the world is “small” when in fact it is not much different from what one would expect by chance, given the degree structure—the distributions of board size and number of directorships held—observed in the social network of corporate governance.

3 Results

3.1 Data

To investigate the social network of corporate boards we use data from the United States, the United Kingdom and Germany. Since these economies have considerably different corporate governance structures they provide an interesting test of the robustness of the random graph model developed in Section 2.

Basic data for each economy is presented in Table 2. The United States data are kindly supplied by the *Corporate Library*. The data set are a snapshot (i.e. a cross section) of publicly traded firms at February 2003. Table 2 illustrates that there are 13,330 unique directors occupying 17,277 separate director seats at 1,733 firms.⁵ This data set is larger than that used by previ-

⁵As directors can be members of more than one board there are fewer unique directors than board positions.

ous US small-world studies of boards of directors e.g. Davis et al. (2003). The average size of a US board is about 10 members and each director, on average, has 1.63 directorships (including the directorship at his or her main company). An individual occupying only one board position is a one-board director. Analogously, a person with two directorships is a two-board director. In the USA the overwhelming majority of directors (about 80%) have only one directorship. About 13% hold precisely two positions, implying that a very small fraction of directors (7%) hold more than two director posts.

The United Kingdom data are supplied by *Hemmington Scott*. The data are a snap shot of the population of publicly traded firms at March 2002. The data has 11,541 directors sitting on 2236 boards. This represents 14,552 director seats. The average board size is 6.51.⁶ Approximately, 84% of directors are one-board directors and 10% are two-board directors. In the USA and the UK around 95% of board members hold either one or two board positions: only a distinct minority board members hold any more than this.

The German data are supplied by *Bureau van Dijk* electronic publishing. Specifically, Germany was selected from the *AMADEUS* database.⁷ The data are a snap shot of publicly traded firms whose fiscal year ended in either 2001 or 2002. The German boards of directors in our analysis consist of the combined lists of both management and supervisory boards. For Germany we selected companies with 300 or more employees and whose board consisted of at least four members. Both publicly quoted firms (*Aktiengesellschaft*, or *AG*) and private limited liability firms (*Gesellschaft mit beschränkter Haftung*, or *GmbH*) are therefore included in the selection criteria. The sample consists of 2,354 firms with 12,747 unique directors who occupy 14,904 director seats. The average board size is 6.33. The average number of directorships per person is 1.45.⁸

3.2 The small-world of corporate boards

Our empirical results are presented in Table 3. For each projection (board and director) we report the number of vertices, the degree, the path lengths, and the clustering coefficients. We present the expected values arising from the theory contained in Section 2 and actual values from the data.

The principal finding arising from Table 3 is the remarkable alignment between the theoretical predictions and actual data. The random graph model of Section 2, based on the known degree distributions of the bipartite representation, predicts almost perfectly the actual data. This conclusion is valid for the

⁶Prior UK research suggests boards are a bit larger than this, but our data includes even the smallest of boards—a population typically excluded from previous studies. In our data board size is an increasing function of firm size measured by market capitalization.

⁷The acronym AMADEUS stands for: Analyse MAJOR Databases from EUROpean Sources

⁸Note the number of vertices in the largest connected component, expressed as a percentage of available vertices, is much smaller in Germany compared to the Anglo-American data. For example, the largest component for US boards is approximately 85% (i.e. 1,473 boards from 1,733) but for Germany it is only 25% (i.e. 582 boards from 2,354). This may occur if the worlds of the AG and the GmbH don't overlap much.

Table 2: *Boards of Directors in the United States, United Kingdom and Germany.*

	USA	UK	Germany
Director seats	17,277	14,552	14,904
Number of unique directors	13,330	11,541	12,747
Number of firms	1,733	2,236	2,354
Average board size	9.97	6.51	6.33
Average number of directorships	1.63	1.84	1.45
One-board director (percent)	80.37%	84.25%	88.33%
Two-board director (percent)	13.02%	10.08%	8.92%
Connected component - firms	1473	1732	582
Connected component - directors	11,057	8,850	4,185

United States, the United Kingdom and Germany despite their quite different corporate governance structures.

First consider the expected degree distribution. For the United States the director projection indicates that a given director is expected to be connected to about 13 directors – and indeed this is what arises in the data (i.e. about 13.46). Also, a given board is expected to be connected to about eight other boards – and again the actually observed data gives rise to a figure of about seven boards. This close correspondence between theory and actual data is also observed for the United Kingdom and Germany.

The small worlds discussed in Section 2 are characterized by relatively short path lengths and high clustering. Table 3 provides evidence on this. For the United States the expected path length in the director projection is about 4.2 (i.e. the expected number of steps it takes for a director to reach any other director is about four). The actual value is only slightly larger at about 5.1. Looking at the board projection for the United States we reach a similar conclusion – theoretical and actual path lengths are broadly similar. If one considers the corresponding path lengths for the United Kingdom we also observe much agreement between the theoretical expectation and the actual data counterpart. Overall, we conclude that the observed path lengths for corporate boards in the United States, the United Kingdom and Germany are no different from what we would expect from a random graph with a known degree distribution. In this sense the world of corporate directors is certainly “small” but no smaller than one would expect theoretically.

However, we should stress that the short-path length outcome is still quite a remarkable finding. For any of the economies considered here the mean geodesic is a very small fraction of the number of vertices in the large connected component. For instance, the mean geodesic (i.e., shortest distance) for the US board projection is 4.327 whereas the number of vertices is 1473 – or about 0.3% of the

Table 3: *The Small World of Corporate Directors in the United States, the United Kingdom and Germany.*

Network	Vertices N	Degree		Length		Clustering		
		Expected $\langle z \rangle$	Actual z_{Actual}	Expected $\langle L \rangle$	Actual L_{Actual}	Expected $\langle C_{\Delta} \rangle$	Actual C_{Actual}	Actual C_{WS}
United States								
Director projection	11057	13.616	13.460	4.228	5.188	0.560	0.556	0.871
Board projection	1473	8.050	7.275	3.510	4.327	0.163	0.167	0.225
United Kingdom								
Director projection	8,850	9.069	8.981	4.795	6.462	0.546	0.612	0.889
Board projection	1732	6.073	5.709	4.116	5.579	0.327	0.376	0.376
Germany								
Director projection	4,185	15.103	14.546	3.504	6.398	0.622	0.719	0.926
Board projection	582	7.337	4.553	2.947	6.108	0.318	0.577	0.413

vertices in the largest connected component. Such short path lengths can act as important and powerful routes for the spread of business practices, information or rumor.⁹

The results for the clustering coefficients contained in Table 3 also show remarkable agreement between theoretical and actual values. For the United States the actual clustering coefficient for the board projection is 0.167 which can be compared to the theoretically expected value of 0.163. Again, the amount of clustering is not much different from what would be expected to occur in a random graph, with known degree distributions. Accordingly, the world of US boards and directors is no more “clubby” than would be expected by chance. Table 3 shows that there is tendency for British directors to be slightly more “clubby” than would be expected than by chance (0.612 versus 0.546) and for German Boards to be more clustered than would be expected by chance (0.577 versus 0.318). For each economy, we also present the Watts-Strogatz clustering measure but from our theoretical analysis we have no way of knowing if this diverges significantly from what would be expected. It is reported since it has been used in the prior literature.

We conclude this subsection by re-stating our main finding: the aggregate network of corporate boards and directors is actually no smaller than would be expected by chance from a random graph model based on the observed distributions of board and director degree. That is, it is not much smaller than one would expect once one has taken into account the very broadest features—the distributions of sizes of board size and of working habits (number of directorships held)—of the world of corporate governance.¹⁰

3.3 Degree correlation

The generating function approach also allows us to compute the entire theoretical degree distribution: not simply the average degree z recorded in Table 3. We can thus plot both the theoretical degree distribution and the one we actually observe and see whether they too are in agreement: Figure 4 contains the results. The circles connected by dashes are observed frequencies while the solid line is the prediction from the generating function. The other dashed (no-circles) curve is the prediction derived from a Poisson distribution that has the same mean degree z as the data.

Focusing on the director results we see, once again, there is remarkable—near perfect—coincidence between the theoretical distribution and the actual data. The random graph model based on empirical degree distributions is a good predictor of the real world social network of directors. In contrast, the

⁹For a Poisson random graph the mean geodesic can be expressed as $l = \log(N)/\log(z)$. So, in the case of the US board projection this is approximately $l = \log(1473)/\log(8.98) = 3.324$. The Poisson random graph, like the random graph with a constrained degree distribution, predicts a short mean geodesic relative to the number of vertices.

¹⁰The qualitative conclusions reached here are unchanged when we investigated alternative data sets. For instance, we studied the 100 largest firms ranked by market capitalization in the United States and the United Kingdom; Executive and Non-executive directors separately in the United Kingdom; and the set of publicly quoted (*AG*) firms only in Germany.

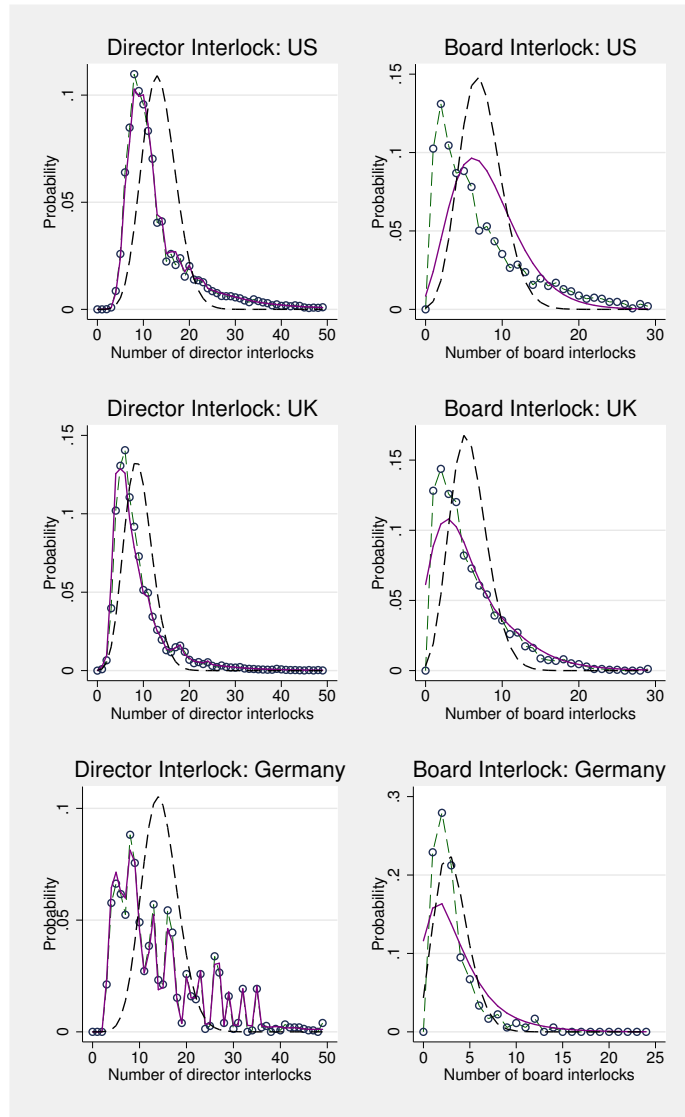


Figure 4: *Board and director interlocks in the United States, United Kingdom and Germany. The left-hand side of the Figure are the degree distributions arising in the director projection. The right-hand side are those arising in the board projection.*



Figure 5: *Quantile and OLS regressions for the Colleague connectivity model. Colleague connectivity is the average number of boards one’s colleagues sit on. Director interlock is the number of boards you sit on. The x-axis is a quantile scale; the y-axis is the estimated coefficient from a regression of colleague connectivity on director interlock*

Poisson random graph distribution (dashed line) with the same mean as the data clearly does not give rise to the actual data.

The board interlock data, however, are much more interesting and revealing. Once again, the Poisson random graph distribution does not fit the data at all well. But, the predictions from the random graph model do not explain perfectly the entire distribution of real-world board interlocks either. Our random graph model under-predicts the frequency of firms with relatively low and relatively high numbers of interlocking boards. Also, the theory over-predicts the frequency of firms with moderate numbers of interlocks. This result, first noted by Newman et al. (2001) for a small US data set, is here verified for comprehensive surveys of the United States, the United Kingdom and for Germany. One should consider a random graph with appropriately constrained degree distributions as the baseline model about which our expectations regarding boards of directors network structure should be evaluated Newman et al. (2001). So, the next question to ask is: why does this anomaly arises?

One plausible explanation is that a director who sits on few boards has colleagues who also sit on few boards. Analogously, a director who sits on many boards has colleagues who are also members of many boards. In short, big-hitters associate with big-hitters and low-hitters associate with other low-hitters. Sociologists refer to this idea homophily—the tendency for like to associate with like (McPherson, Smith-Lovin & Cook (2001)). The homophily effect gives rise

to positive degree correlation. As McPherson et al. (2001) remark “Similarity breeds connection.”

We tested for positive degree correlation (homophily) in the data using quantile and ordinary least squares regression methods (see Koenker & Hallock (2001) and Koenker & Bassett (1978)). The results are contained in Figure 5. We define the outcome variable as the average number of directorships held by a director’s colleagues at the director’s own board. For short-hand we will refer to this variable as *colleague connectivity*. We define a covariate as the number of boards a director sits on. As short-hand we refer to this as *director interlock*. We then performed a regression of *colleague connectivity* on *director interlock* with the expectation of a positive sign.

For the coefficient on the *director interlock* covariate we plot the 19 separate quantile regression estimates from 0.05 to 0.95 rising in equal steps of 0.05. This is the solid line plot in Figure 5. These point estimates are interpreted as the impact of a one unit change in the number of *director interlocks* on the *colleague connectivity* variable. In Figure 5 the horizontal axis is the quantile scale and the vertical axis is the covariate effect on *colleague connectivity*. The dot-dash lines are the 95% confidence intervals for the quantile estimates. The horizontal dashed line is the ordinary least squares estimate.

The evidence contained in Figure 5 broadly supports the idea of positive degree correlation, or homophily. For the United States, the United Kingdom and Germany the OLS estimate is positive and significant. Directors who sit on many boards apparently do so in the company of others who also sit on many boards. However, this ordinary least square estimate conceals important information about the effect of *director interlocks*. In the United States, at the 5th percentile of the conditional distribution, the effect of a director interlock on one’s colleagues propensity to have many directorships is about 0.1. This estimate increases in magnitude towards the center of the distribution before falling back. Similar, effects are observed for the United Kingdom and Germany. Note that the quantile estimates are all positive: we conclude that there is positive degree correlation (homophily) in these data sets.

4 Concluding remarks

In this paper we have presented a random graph model to explain the social network structure of boards of directors. Our paper is informed by developments in graph theory and the structure of complex networks – Newman (2003). We have investigated the conditions under which the social network of corporate directors can be described as a “small-world” and the related question of how “small” we expect that world to be. We have shown how a random graph with a known degree distribution gives rise to an expected level of *clustering* determined by the structure of the actual network. We also illustrated that the theoretical degree distribution can be determined using generating function machinery.

Our empirical results show that the random graph model is remarkably

good at explaining board structure and connectedness in the United States, the United Kingdom and Germany. Although there are small-world traits such as “clustering” and “short-paths” in the corporate world, they are no more pronounced than would be expected by chance in our random graph model. This is not to suggest that boards of directors really are assembled at random, only to say that when the social network of corporate governance is viewed in the large—and especially when it is viewed through the lens of small-world statistics—it reveals no more systematic structure, no stronger tendency toward “smallness”, than one should expect find by chance.

Alternatively one could say that small-world measures, at least when applied to the entirety of a country’s corporate world, are rather less informative—in the sense that they reveal less structure—than one might have hoped. An important exception appears in the detailed comparison of real and predicted distributions of board interlocks. There we find good evidence for positive degree correlation: directors who sit on many boards appear to do so in the company of others who also sit on many boards. This result may help to explain the distribution of boards interlocks, but one would need to do further analysis to establish this.

The three data sets discussed here are among the most comprehensive yet studied from the small-world viewpoint of directors, but nonetheless they leave many of the most interesting questions tantalizingly beyond reach. For example, all our data sets are snapshots of their respective corporate networks: it would be very interesting to collect similar data over a period of years (ideally lasting many times longer than a typical director’s term of service) in order to look at the dynamics of connections within the corporate world. Further, the analysis in this paper is purely cartographic: we have mapped out structural features of the corporate world, but have yet to exploit this understanding as, for example, an aid to studying or modelling the flow of information or business practice. This exercise we leave for future research.

References

- Alon, N. & J.H. Spencer (2000), *The Probabilistic Method*, second edn, John Wiley & Sons, New York.
- Baum, Joel A. C., Andrew V. Shiplov & Tim J. Rowley (2003), ‘Where do small worlds come from?’, *Industrial and Corporate Change* **12**, 697–725.
- Bollobás, B. (2001), *Random Graphs*, second edn, Cambridge University Press, Cambridge.
- Booth, J.R. & D.N. Deli (1996), ‘Factors affecting the number of outside directorships held by CEOs’, *Journal of Financial Economics* **40**, 81–104.
- Conyon, Martin J. & Joachim Schwalbach (2000), ‘Executive compensation: Evidence from the UK and Germany’, *Long Range Planning* **33**, 504–526.
- Conyon, Martin J. & Kevin J. Murphy (2000), ‘The prince and the pauper? CEO pay in the United States and United kingdom’, *Economic Journal* **110**, 640–671.
- Core, John, Robert Holthausen & David Larcker (1999), ‘Corporate governance, chief executive officer compensation and firm performance’, *Journal of Financial Economics* **51**, 371–406.
- Cormen, T.H., C.E. Leiserson & R.L. Rivest (2001), *Introduction to Algorithms*, second edn, MIT Press, Cambridge MA.
- Dahya, Jay, John J. McConnell & Nickolaos G. Travlos (2002), ‘The increasing use of non-executive directors: Its impact on UK board structure and governance arrangements’, *Journal of Finance* **57**, 461–494.
- Davis, Gerald F. & Henrich R. Greve (1997), ‘Corporate elite networks and governance changes in the 1980s’, *American Journal of Sociology* **103**, 1–37.
- Davis, Gerald F., Mina Yoo & Wayne E. Baker (2003), ‘The small world of the corporate elite, 1982-2001’, *Strategic Organization* pp. 301–326.
- Dooley, Peter C. (1969), ‘The interlocking directorate’, *American Economic Review* **59**, 314–324.
- Erdős, P. & A. Rényi (1959), ‘On random graphs’, *Publ. Math. Debrecen* **6**, 290–297.
- Erdős, P. & A. Rényi (1960), ‘On the evolution of random graphs’, *Publ. Math. Inst. Hung. Acad. Sci., Ser. A* **5**, 17–61.
- Fama, Eugene F. & Michael C. Jensen (1983), ‘Separation of ownership and control’, *Journal of Law and Economics* **26**, 301–325.

- Franks, Julian & Colin Mayer (2001), 'Ownership and control of German corporations', *The Review of Financial Studies* **14**, 943–977.
- Hallock, Kevin F. (1997), 'Reciprocally interlocking board of directors and executive compensation', *Journal of Financial and Quantitative Analysis* **32**, 331–344.
- Hermalin, Benjamin & Michael Weisbach (2003), 'Boards of directors as an endogenously determined institution: a survey of the economic literature', *FRBNY Economic Policy Review* **9**, 7–22.
- Koenker, Roger & Gilbert Bassett (1978), 'Regression quantiles', *Econometrica* **15**, 33–50.
- Koenker, Roger & Kevin Hallock (2001), 'Quantile regression', *Journal of Economic Perspectives* **46**, 143–156.
- Kogut, Bruce & Gordon Walker (2001), 'The small world of Germany and the durability of national networks', *American Sociological Review* **66**, 317–335.
- McPherson, Miller, Lynn Smith-Lovin & James Cook (2001), 'Birds of a feather: Homophily in social networks', *Annual Review of Sociology* **27**, 415–444.
- Newman, Mark E. J. (2003), 'The structure and function of complex networks', *SIAM Review* **45**, 167–256.
- Newman, Mark E.J., Duncan J. Watts & Steven H. Strogatz (2002), 'Random graph models for social networks', *Proc. Nat. Acad. Sci. U.S.A.* **99**, 2566–2572.
- Newman, Mark E.J., Steven H. Strogatz & Duncan J. Watts (2001), 'Random graphs with arbitrary degree distributions and their applications', *Physical Review E* **64**.
- Peasnell, Ken, Peter Pope & Steven Young (2003), 'Managerial equity ownership and the demand for outside directors', *European Financial Management* **9**, 99–118.
- Schilling, Melissa A. & C. Phelps (2004), Interfirm collaboration networks: the impact of network structure on rates of innovation. Stern School Working paper, New York University.
- Wasserman, S. & K. Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.
- Watts, Duncan J. (1999), *Small Worlds: The Dynamics of Networks between Order and Randomness*, Cambridge University Press, Princeton Univ. Press, Princeton N.J.
- Wilf, Herbert S. (1990), *Generating functionology*, Academic Press Professional, Inc., Academic Press Professional, Inc. San Diego, CA, USA.

Young, Steven (2000), 'The Cadbury committee, corporate performance, and top management turnover', *Journal of Business Finance and Accounting* **27**, 1311–1342.