J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

arXiv:0902.0878v1 [q-fin.GN] 5 Feb 2009

The backbone of complex networks of corporations: Who is controlling whom?

J.B. Glattfelder, S. Battiston

Chair of Systems Design, ETH Zurich, Kreuzplatz 5, 8032 Zurich, Switzerland

`jglattfelder@ethz.ch`, `sbattiston@ethz.ch`

Abstract

We present a methodology to extract the backbone of complex networks in which the weight and direction of links, as well as non-topological state variables associated with nodes play a crucial role. This methodology can be applied in general to networks in which mass or energy is flowing along the links. In this paper, we show how the procedure enables us to address important questions in economics, namely how control and wealth is structured and concentrated across national markets. We report on the first cross-country investigation of ownership networks in the stock markets of 48 countries around the world. On the one hand, our analysis confirms results expected on the basis of the literature on corporate control, namely that in Anglo-Saxon countries control tends to be dispersed among numerous shareholders. On the other hand, it also reveals that in the same countries, control is found to be highly concentrated at the global level, namely lying in the hands of very few important shareholders. This result has previously not been reported, as it is not observable without the kind of network analysis developed here.

PACS numbers: 89.65.Gh, 02.50.-r, 05.45.Df, 64.60.aq

1   Introduction

The empirical analysis of real-world complex networks has revealed unsuspected regularities such as scaling laws which are robust across many domains, ranging from biology or computer systems to society and economics [1, 2, 3, 4]. This has suggested that universal or at least generic mechanisms are at work in the formation of many such networks. Tools and concepts from statistical physics have been crucial for the achievement of these findings [5, 6].

In the last years, in order to offer useful insights into more detailed research questions, several studies have started taking into account the specific meaning of the nodes and links in the various domains the the real-world networks pertain to [7, 8]. Three levels of analysis are possible. The lowest level corresponds to a purely topological approach (best epitomized by a binary adjacency matrix, where links simply exists or do not). Allowing the links to carry weights [7], or weights and direction [9], defines the second level. Only recent studies have started focusing on the third level of detail, in which the nodes themselves are assigned a degree of freedom, sometimes also called fitness. This is a non-topological state variables which shapes the topology of the network [8, 10, 11].

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Indeed, when analyzing real-world networks, considering all three levels can yield new insights which would otherwise remain unobserved. For instance, in the present paper, the identification of the key players in the networks under study is only possible if the network analysis takes into account a non-topological variable (namely, the value of the market capitalization of the listed companies). In doing so, we are able to show that in markets where the control of corporations tends to be more evenly distributed across many shareholders, unexpectedly, the control, from a global point of view, tends to be more concentrated in the hands of few shareholders. This result is in contrast with previously held views in the economics literature.

However, considering all three levels of detail does not guarantee per se that new insights can be gained. It is also essential that the standard measures utilized in the analysis of complex networks are appropriately adapted to the specific nature of the network under investigation. For instance, the study of the degree distribution in various real-world networks has revealed universal features across different domains [12]. In many cases however, the degree of the nodes is not a suitable measure of connectivity [7, 10]. In this paper, we introduce novel quantities, analogous to in- and out-degree, which are better suited for networks in which the relative weight of the links are important.

The physics literature on complex economic networks has previously focused on boards of directors [13, 14], market investments [10, 15], stock price correlations [16, 17] and international trade [18, 19]. In this context, the present work represents the first comprehensive cross-country analysis of 48 stock markets world-wide. The paper introduces a novel algorithm able to identify and extract the backbone in the networks of ownership relations among firms. Notably, we also provide a generalization of the method applicable to networks in which weights and direction of links, as well as non-topological state variables assigned to the nodes play a role. In particular, the method is relevant for networks in which there is a flow of mass (or energy) along the links and one is interested in identifying the subset of nodes where a given fraction of the mass of the system is flowing.

In this paper, we show how this type of complex network analysis can address research questions that are important in economics. To this aim, we need to briefly review the relevant literature. In economics, the corporate finance and corporate governance literature addresses issues related to the notions of ownership and control. As an example, the question to what extent the economic activities of a country are in the control of one or more groups of few actors has been a recurring theme. The answer has important implications in terms of competition, innovation, and even for political power [20].

There is a vast body of literature on corporate control that focuses on corporations as individual units. The research topics this field of study addresses can be grouped into three major categories. Firstly, analyzing the dispersion or concentration of control [21, 22, 23]. Secondly, empirically investigating how the patterns of control vary across countries and what determines them [24, 25].

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

And thirdly, studying the impact of frequently observed complex ownership patterns [26, 27, 28, 29] such as so-called pyramids [30] and cross-shareholdings (also known as business groups) [31].

In addition, research in cooperative game theory analyzing political voting games has resulted in the development of so-called power indices [32, 33]. These ideas have been applied to coalitions of shareholders voting at Shareholders Meetings [34].

It should be noted that most previous empirical studies did not build on the idea that ownership and control define a vast complex network of dependencies. Instead, they selected samples of specific companies and looked only at their local web of interconnections. These approaches are unable to discern control at a global level. This emphasizes the fact that the bird's-eye-view given by a network perspective is important for unveiling overarching relationships. Remarkably, the investigation of the financial architecture of corporations in national or global economies taken *as a whole* is just at the beginning [10, 35, 36].

In a nutshell, the research questions arising from the analysis of ownership networks can be summarized as follows: what is the map of corporate control? This entails the study of the global distribution of control next to identifying the degree of fragmentation or integration of such control structures. These questions can be posed at a *country* or at a *world-wide* level.

In this paper, we focus only on the issue of how control is distributed, at the country level, based on the knowledge of the ownership ties. Indeed, although control is exercised in many subtle ways, ownership is certainly one of the main vehicles of control. As mentioned, our aim is to investigate the nature of ownership networks, that is, a web of shareholding relations of quoted companies and their shareholders in 48 country's stock markets. In detail, we address the issue of how control and wealth is structured in these markets. As a first step, we propose a new model to estimate corporate control based on the knowledge of the ownership ties. We then not only incorporate all three levels of network analysis, but also consider higher orders of neighborhood relations, next to accounting for all indirect ownership ties in our study. In this respect, to our knowledge, there exists no comparable work of this kind in the literature. Our methodology allows us to identify and extract the core subnetwork where most of the value of the stock market resides, called the *backbone of control*. The analysis of these structures reveals previously unobservable results. Not only is the local dispersion of control associated with a global concentration of control and value, in addition, the local concentration of control is related to a global dispersion of control and value. In detail, an even distribution of control at the level of individual corporations (typical of Anglo-Saxon markets) is accompanied by a high concentration of control and value at the global level. This novel observation means that, in such countries, although stocks tend to be held by many shareholders, the market as a whole is actually controlled by very few shareholders. On the other hand, in countries where the control is locally concentrated (e.g., European states), control and value is dispersed at the global level, meaning that there is a large number of shareholders controlling few corporations.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

The paper is organized as follows. Sec. 2 describes the dataset we used. In Sec. 3 we introduce and discuss our methodology and perform a preliminary topological analysis of the networks. Sec. 4 describes the backbone extraction algorithm. In particular, we show that the method can be generalized by providing a recipe for generic weighted and directed networks. The section also introduces classification measures which are employed for the backbone analysis in Sec. 5. Finally, Sec. 6 summarizes our results and concludes the paper.

2 The Dataset

We are able to employ a unique dataset consisting of financial data on public companies and their shareholders in global stock markets. We constrain our analysis to a subset of 48 countries: United Arab Emirates (AE), Argentina (AR), Austria (AT), Australia (AU), Belgium (BE), Bermuda (BM), Canada (CA), Switzerland (CH), Chile (CL), China (CN), Germany (DE), Denmark (DK), Spain (ES), Finland (FI), France (FR), United Kingdom (GB), Greece (GR), Hong Kong (HK), Indonesia (ID), Ireland (IE), Israel (IL), India (IN), Iceland (IS), Italy (IT), Jordan (JO), Japan (JP), South Korea (KR), Kuwait (KW), Cayman Islands (KY), Luxembourg (LU), Mexico (MX), Malaysia (MY), Netherlands (NL), Norway (NO), New Zealand (NZ), Oman (OM), Philippines (PH), Portugal (PT), Saudi Arabia (SA), Sweden (SE), Singapore (SG), Thailand (TH), Tunisia (TN), Turkey (TR), Taiwan (TW), USA (US), Virgin Islands (VG), South Africa (ZA). In the following, the countries will be identified by their two letter ISO 3166-1 alpha-2 codes given in the parenthesis above. To assemble the ownership networks of the individual countries, we select the stocks in the country's market and all their available shareholders, who can be natural persons, national or international corporations themselves, or other legal entities.

The data is compiled from Bureau van Dijk's ORBIS database[1]. In total, we analyze 24877 corporations (or stocks) and 106141 shareholding entities who cannot be owned themselves (individuals, families, cooperative societies, registered associations, foundations, public authorities, etc.). Note that because the corporations can also appear as shareholders, the network does not display a bipartite structure. The stocks are connected through 545896 ownership ties to their shareholders. The database represents a snapshot of the ownership relations at the beginning of 2007. The values for the market capitalization, which is defined as the number of outstanding shares times the firm's market price, are also from early 2007. These values will be our proxy for the size of corporations and hence serve as the non-topological state variables.

We ensure that every node in the network is a distinct entity. In addition, as theoretically the sum of the shareholdings of a company should be 100%, we normalize the ownership percentages if the sum is smaller due to unreported shareholdings. Such missing ownership data is nearly always due to their percentage values being very small and hence negligible.

---

[1]http://www.bvdep.com/orbis.html.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

3   A 3-Level Network Analysis

Standard network analysis focuses on topics like degree distribution, assortativity, clustering coefficients, average path lengths, connected components, etc. However, our specific interest in the structure of control renders most of these quantities inappropriate.

For instance, the out-degree measures, in an ownership network, the number of firms in which a shareholder has invested. A high out-degree does not imply high control since the shares could be very small. Similarly, the in-degree, revealing the number of shareholders a corporation has, gives little insight into the amount of influence these shareholders can exert. In Sec. 3.2 we therefore extend the notion of degree to fit our context. Consequently, it is also not clear how to interpret degree-degree correlations, i.e., (dis-) assortativity.

The clustering coefficient defined for undirected graphs is equivalent to counting the number of triangles in a network. It does not have an obvious interpretation in the directed case, since an undirected triangle can correspond to several directed triangle configurations. Clustering coefficients have been introduced for weighted and undirected networks [7], next to weighted and directed networks [37]. However, these definitions only consider paths of length two. In contrast, in this paper, we use a measure of control that consider all paths of all lengths (see Sec. 3.5). Indeed, the knowledge of all the stocks reachable from any particular shareholder represents nothing else than a definition of indirect control.

For similar reasons, the average path length for the undirected graph does not have an interpretation in terms of control. Therefore, for our purposes, it also does not make sense to compute the small-world property (which is based on the two previously discussed quantities) of these real-world networks.

On the other hand, an analysis of the connected components may provide insights into the degree of fragmentation of the capital markets and we briefly address this issue in the following section. We then introduce extensions of existing network measures and define new quantities that better suit the ownership networks which are subsequently analyzed at all three levels of resolution in Sec. 4.

3.1   Level 1: Topological analysis

The network of ownership relations in a country is very intricate and a cross-country analysis of some basic properties of these networks reveals a great level of variability.

For example, an analysis of the number and sizes of connected components unveils a spectrum ranging from a single connected component in IS to 459 in the US. With a size of 18468, the largest connected component in the US is bigger than any single national ownership network in our sample.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
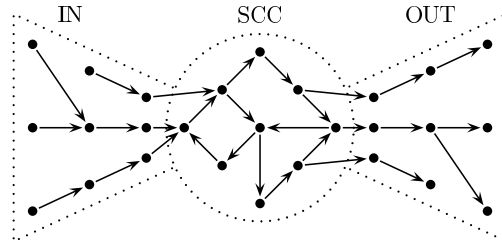Who is controlling whom?

Figure 1: Schematic illustration of a bow-tie topology: the central area is the strongly connected component (SCC), where there is a path from each node to every other node, and the left (IN) and right (OUT) sections contain the incoming and outgoing nodes, respectively.

Many small components correspond to a fragmented capital market while a giant and dense component corresponds to an integrated market. It is however not very clear what such connected components reveal about the structure and distribution of control. The same pattern of connected components can feature many different configurations of control. Therefore, it makes sense to move on to the next level of analysis by introducing the notion of direction. Now it is possible to identify strongly connected components. In terms of ownership networks, these patterns correspond to sets of corporations where every firm is connected to every other firm via a path of indirect ownership. Furthermore, these components may form bow-tie structures, akin to the topology of the World Wide Web [38]. Fig. 1 illustrates an idealized bow-tie topology. This structure reflects the flow of control, as every shareholder in the IN section exerts control and all corporations in the OUT section are controlled.

We find that roughly two thirds of the countries' ownership networks contain bow-tie structures (see also [39]). Indeed, already at this level of analysis, previously observed patterns can be rediscovered. As an example, the countries with the highest occurrence of (small) bow-tie structures are KR and TW, and to a lesser degree JP. A possible determinant is the well known existence of so-called business groups in these countries (e.g., the *keiretsu* in JP, and the *chaebol* in KR) forming a tightly-knit web of cross-shareholdings (see the introduction and references in [31] and [40]). For AU, CA, GB and US we observe very few bow-tie structures of which the largest ones however contain hundreds to thousands of corporations. It is an open question if the emergence of these mega-structures in the Anglo-Saxon countries is due to their unique "type" of capitalism (the so-called Atlantic or stock market capitalism, see the introduction and references in [41]), and whether this finding contradicts the assumption that these markets are characterized by the absence of business groups [31].

Continuing with this line of research would lead to the question of how control is fragmented (e.g., investigations of the distribution of cluster sizes, cluster densities, etc.). Further analyzing this issue at the third level would require the weight of links and non-topological variables of the nodes to be considered as well. As our current interest is devoted to the first question of how
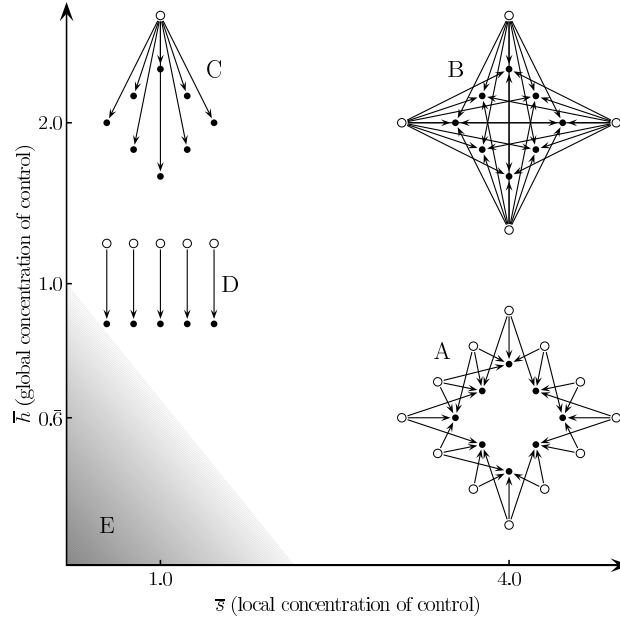
J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Figure 2: The map of control: illustration of idealized network topologies in terms of local dispersion of control (x-axis) vs. global concentration of control (y-axis); shareholders and stocks are shown as empty and filled bullets, respectively; arrows represent ownership; consult the discussion in the text; $\overline{s}$ and $\overline{h}$ will be introduced in Sec. 4.4; see Fig. 11 for the empirical results.

control is distributed, we do not further investigate the nature of the connected components. We ask instead what structures can be identified that reflect the concentration of control. Our proposed methodology answers this question by extracting the core structures of the ownership networks — the backbones — unveiling the seat of power in national stock markets (see Sec. 4).

Fig. 2 anticipates the possible generic backbone configurations resulting from local and global distributions of control. Moving to the right-hand side of the x-axis the stocks have many shareholders (local dispersion of control), whereas stocks on the very left side have only one shareholder each. The y-axis depicts the global concentration of control, i.e., how many shareholders are controlling all the stocks in the market. Moving up the y-axis, the stocks are held by fewer and fewer shareholders. There is a consistency constraint on the coordinates that are allowed and region (E) is excluded. Possible network configurations are (A) many owners sharing many stocks, (B) few shareholders holding many stocks, (C) a single shareholder controlling all the stocks and (D) a situation with an equal number of shareholders, ownership ties and stocks. Note that (A) does not necessarily need to be a connected structure as many fragmented network configurations can result in such coordinates.
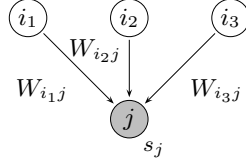
J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Figure 3: Definition of the concentration index s_j, measuring the number of prominent incoming edges, respectively the effective number of shareholders of the stock j. When all the weights are equal, then $s_j \quad = \quad k_j^{in}$, where k_j^{in} is the in-degree of vertex j. When one weight is overwhelmingly larger than the others, the concentration index approaches the value one, meaning that there exists a single dominant shareholder of j.

3.2 Level 2: Extending the notions of degree

In graph theory, the number of edges per vertex i is called the *connectivity degree* and is denoted by k_i. If the edges are oriented, one has to distinguish between the in-degree and out-degree, k^{in} and k^{out}, respectively. When the edges are weighted, the corresponding quantity is called *strength* [7]:

$$k_i^w := \sum_j W_{ij}. \tag{1}$$

Note that for weighted and oriented networks, one has to distinguish between the in- and out-strengths, k^{in-w} and k^{out-w}, respectively.

However, the interpretation of $k^{in/out-w}$ is not always straightforward for real-world networks. In the case of ownership networks, as mentioned in Sec. 3, there is no useful meaning associated with these values. In order to provide a more refined and appropriate description of weighted ownership networks, we introduce two quantities that extend the notions of degree and strength in a sensible way.

The first quantity to be considered reflects the relative importance of the neighbors of a vertex. More specifically, given a vertex j and its incoming edges, we focus on the originating vertices of such edges, as shown in Fig. 3. The idea is to define a quantity that captures the relative importance of incoming edges.

When there are no weights associated with the edges, we expect all edges to count the same. If weights have a large variance, some edges will be more important than others. A way of measuring the number of prominent incoming edges is to define the *concentration index* as follows:

$$s_j := \frac{\left(\sum_{i=1}^{k_j^{in}} W_{ij} \right)^2}{\sum_{i=1}^{k_j^{in}} W_{ij}^2}. \tag{2}$$

J.B. Glattfelder, S. Battiston:
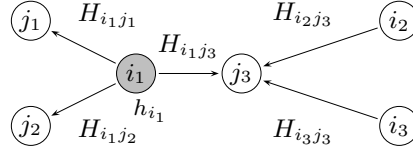The backbone of complex networks of corporations:
Who is controlling whom?

Figure 4: The definition of the control index h_i, measuring the number of prominent outgoing edges. In the context of ownership networks this value represents the effective number of stocks that are controlled by shareholder i. Note that to obtain such a measure, we have to consider the fraction of control H_{ij}, which is a model of how ownership can be mapped to control (see the discussion in Sec. 3.5).

Note that this quantity is akin to the inverse of the Herfindahl index extensively used in economics as a standard indicator of market concentration [42]. Indeed, already in the 1980s the Herfindahl index was also introduced to measure ownership concentration [43]. Notably, a similar measure has also been used in statistical physics as an order parameter [44]. In the context of ownership networks, s_j is interpreted as the effective number of shareholders of the stock j. Thus it can be interpreted as a measure of control from the point of view of a stock.

The second quantity to be introduced measures the number of important outgoing edges of the vertices. For a given vertex i, with a destination vertex j, we first define a measure which reflects the importance of i with respect to all vertices connecting to j:

$$H_{ij} := \frac{W_{ij}^2}{\sum_{l=1}^{k_j^{in}} W_{lj}^2}. \tag{3}$$

This quantity has values in the interval $(0, 1]$. For instance, if $H_{ij} \approx 1$ then i is by far the most important destination vertex for the vertex j. For our ownership network, H_{ij} represents the *fraction of control* shareholder i has on the company j. For an interpretation of H_{ij} from an economics point of view, consult Sec. 3.5.

In a next step, we then define the *control index*:

$$h_i := \sum_{j=1}^{k_i^{out}} H_{ij}. \tag{4}$$

As shown in Fig. 4, this quantity is a way of measuring how important the outgoing edges of a node i are with respect to its neighbors' neighbors. Within the ownership network setting, h_i is interpreted as the effective number of stocks controlled by shareholder i.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

3.3 Distributions of s and h

In this paper, s and h are primarily used in the algorithm that extracts the backbone (see Sec. 4). However, these measures can also provide insights into the patterns of how ownership and control are distributed at a local level.

Fig. 5 shows the probability density function (PDF) of s_j for a selection of nine countries (for the full sample consult [45]). There is a diversity in the shapes and ranges of the distributions to be seen. For instance, the distribution of GB reveals that many companies have more than 20 leading shareholders, whereas in IT few companies are held by more than five significant shareholders. Such country-specific signatures were expected to appear due to the differences in legal and institutional settings (e.g., law enforcement, protection of minority shareholders [25]).

On the other hand, looking at the cumulative distribution function (CDF) of k_i^{out} (shown for three selected countries in the top panel of Fig. 6; the full sample is available at [45]) a more uniform shape is revealed. The distributions range across two to three orders of magnitude. Hence some shareholders can hold up to a couple of thousand stocks, whereas the majority have ownership in less than 10. Considering the CDF of h_i, seen in the middle panel of Fig. 6, one can observe that the curves of h_i display two regimes. This is true for nearly all analyzed countries, with a slight country-dependent variability. Notable exceptions are FI, IS, LU, PT, TN, TW, VG. In order to understand this behavior it is useful to look at the PDF of h_i, shown in the bottom panel of Fig. 6. This uncovers a new systematic feature: the peak at the value of $h_i = 1$ indicates that there are many shareholders in the markets who's only intention is to control one single stock. This observation, however, could also be due to a database artefact as incompleteness of the data may result in many stocks having only one reported shareholder. In order to check that
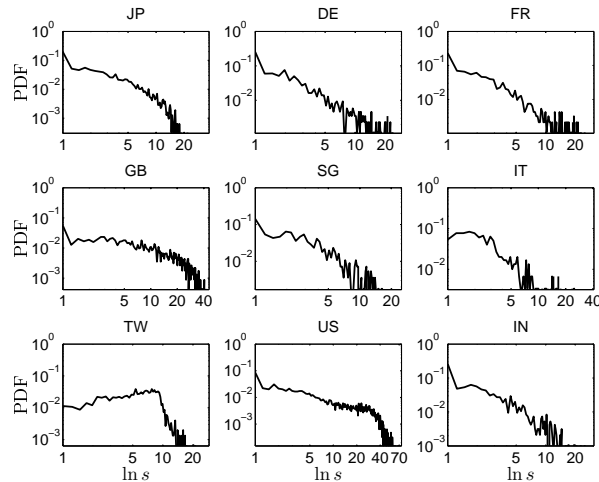

Figure 5: Probability distributions of s_j for selected countries; PDF in log-log scale.
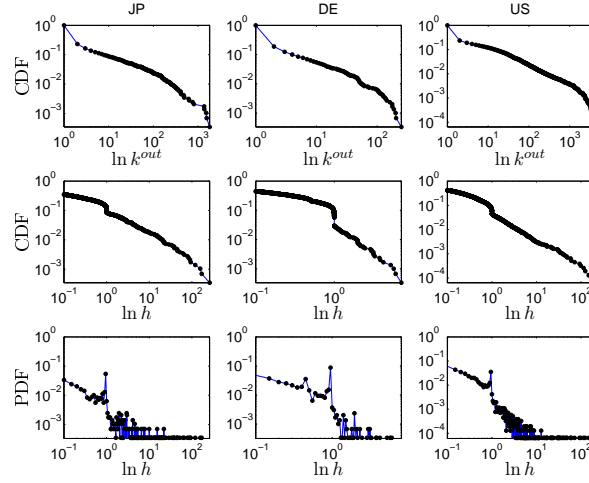
J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Figure 6: Various probability distributions for selected countries: (*top panel*) CDF plot of k_i^{out}; (*middle panel*) CDF plot of h_i; (*bottom panel*) PDF plot of h_i; all plots are in log-log scale.

this result is indeed a feature of the markets, we constrain these ownership relations to the ones being bigger than 50%, reflecting incontestable control. In a subsequent analysis we still observe this pattern in many countries (BM, CA, CH, DE, FR, GB, ID, IN, KY, MY, TH, US, ZA; ES being the most pronounced). In addition, we find many such shareholders to be non-firms, i.e., people, families or legal entities, hardening the evidence for this type of exclusive control. This result emphasizes the utility of the newly defined measures to uncover relevant structures in the real-world ownership networks.

3.4   Level 3: Adding non-topological values

The quantities defined in Eqs. (2) and (4) rely on the direction and weight of the links. However, they do not consider non-topological state variables assigned to the nodes themselves. In our case of ownership networks, a natural choice is to use the market capitalization value of firms in thousand USD, v_j, as a proxy for their sizes. Hence v_j will be utilized as the state variable in the subsequent analysis. In a first step, we address the question of how much wealth the shareholders own, i.e, the value in their portfolios.

As the percentage of ownership given by W_{ij} is a measure of the fraction of outstanding shares i holds in j, and the market capitalization of j is defined by the number of outstanding shares times the market price, the following quantity reflects i's *portfolio value*:

$$p_i := \sum_{j=1}^{k_i^{out}} W_{ij} v_j. \tag{5}$$

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Extending this measure to incorporate the notions of control, we replace W_{ij} in the previous equation with the fraction of control H_{ij}, defined in Eq. (3), yielding the *control value*:

$$c_i := \sum_{j=1}^{k_i^{out}} H_{ij} v_j. \tag{6}$$

A high c_i value is indicative of the possibility to control a portfolio with a big market capitalization value. This newly introduced quantity (extended to also include indirect control relations, as described in the next section) is used in Sec. 4.1 to identify and rank the important shareholders.

3.5 The interpretation and extension of H_{ij}

In Sec. 3.2 the fraction of control, H_{ij}, was introduced from a network perspective as giving the relative importance of node i with respect to all other nodes linking to j. From an economics point of view, it should be emphasized that while ownership is an objective quantity (the percentage of shares owned), control can only be estimated. Several models aiming at deriving control based on the knowledge of ownership have been proposed. In this section we discuss how our new measure overcomes some of the limitations of previous models.

There is a great freedom in how corporations are allowed to map percentages of ownership in their equity capital (also referred to as cash-flow rights) into voting rights assigned to the holders at Shareholders Meetings (e.g., nonvoting shares, dual classes of shares, multiple voting rights, golden shares, voting-right ceilings, etc.). However, empirical studies indicate that in many countries the corporations tend not to exploit all the opportunities allowed by national laws to skew voting rights. Instead, they adopt the so-called one-share-one-vote principle which states that ownership percentages yield identical percentages of voting rights [25, 46].

It is however still not obvious how to compute control from the knowledge of the voting rights. As an example, some simple models introducing a fixed threshold for control have been proposed (with threshold values of 10% and 20% [25] next to a more conservative value of 50% [29]).

Furthermore, *indirect* ownership relations are not negligible. Complex ownership structures themselves can act as vehicles to separate ownership from control. To address the question of how control propagates via indirect ownership, the so-called *integrated model* has been proposed [26]. Consider a sample of n firms connected by cross-shareholdings and pyramidal ownership relations. Let A_{ij}, with $i, j = 1, 2, ..., n$, be the ownership (W_{ij}) or control (H_{ij}) that company i has directly on company j, and $A = [A_{ij}]$ is the matrix of all the links between every one of the n firms. By definition, it holds that

$$\sum_{i=1}^{n} A_{ij} \leq 1; \qquad j = 1, ..., n. \tag{7}$$

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

When some shareholders of company i are not identified or are outside the sample n, the inequality becomes strict. The integrated model accounts for direct and indirect ownership through a recursive computation. The general form of the equation reads

$$\tilde{A}_{ij} := A_{ij} + \sum_n A_{in}\tilde{A}_{nj}, \tag{8}$$

where the tilde denotes integrated ownership or control. This expression can be written in matrix form as

$$\tilde{A} = A + A\tilde{A}, \tag{9}$$

the solution of which is given by

$$\tilde{A} = (I - A)^{-1}A. \tag{10}$$

For the matrix $(I - A)$ to be non-negative and non-singular, a sufficient condition is that the Frobenius root is smaller than one, $\lambda(A) < 1$. This is ensured by the following requirement: in each strongly connected component $\mathcal{S}$ there exists at least one node j such that $\sum_{i\in\mathcal{S}} A_{ij} < 1$. In an economic setting, this means that there exists no subset of k firms ($k = 1,\ldots,n$) that are entirely owned by the k firms themselves. A condition which is always fulfilled in ownership networks [26].

In order to derive the integrated model for the control value defined in Eq. (6), we first solve Eq. (10) for the fraction of control H_{ij} to yield the integrated fraction of control $\tilde{H}_{ij}$, and then sum over the market capitalization of all held assets, v_j, weighted by this value to recover the *integrated control value*:

$$\tilde{c}_i := \sum_{j=1}^{k_i^{out}} \tilde{H}_{ij}v_j. \tag{11}$$

The computation of the fraction of control and the integrated model can be understood in terms of two non-commutative mappings.

There is a further problem in estimating control or power: shareholders do not only act as individuals but can collaborate in shareholding coalitions and give rise to so-called voting blocks. The theory of political voting games in cooperative game theory has been applied to the problem of shareholder voting in the form of so-called *power indices* [47]. However, the employment of power indices for measuring shareholder voting behavior has failed to find widespread acceptance due to computational, inconsistency and conceptual issues [47, 48].

The so-called degree of control, α, was introduced in [43, 49] as a probabilistic voting model measuring the degree of control of a block of large shareholdings as the probability of it attracting majority support in a voting game. Without going into details, the idea is as follows. Consider a shareholder i with ownership W_{ij} in the stock j. Then the control of i depends not only on the value in absolute terms of W_{ij}, but also on how dispersed the remaining shares are (measured

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

by the Herfindahl index). The more they tend to be dispersed, the higher the value of α. So even a shareholder with a small W_{ij} can obtain a high degree of control. The assumptions underlying this probabilistic voting model correspond to those behind the power indices. It is important to realize, that α can only be applied for the largest shareholders, as it gives a minimum cutoff value of 0.5 (even for arbitrarily small shareholdings). As a consequence, computing α for all the shareholders of a company violates Eq. (7) and therefore it cannot be utilized in an integrated model.

Based on the previous discussion, we present a minimal list of requirements a reasonable model of control should fulfil:

1. Define a mapping from $F : (0,1]^N \to (0,1]^N$, for the N shareholding relations $\{W_{ij}\}$, where $F_1(\{W_{ij}\}), \ldots, F_N(\{W_{ij}\})$ represent control and take on continuous values.

2. Be extendable to an integrated version.

3. Sum to one for each stock, as $\sum_j W_{ij}$ in principle does.

4. Emulate the behavior of α for large shareholders.

5. Have an intuitive meaning of controlling power.

6. Be feasible to compute on large networks.

Indeed, our quantity H_{ij} adheres to this small catalogue of requirements. The definition of H_{ij} lies between a linear mapping implied by the one-share-one-vote principle and the fixed-threshold model. It holds that $\sum_j H_{ij} = 1$, for all stocks j. In effect, any shareholder gaining control will be offset by shareholders loosing control. For large shareholders, the analytical expressions of H_{ij} and α share very similar behavior (a detailed discussion of this point is beyond the scope of this paper). This means that to some extent our measure of control can take possible strategic alliances of shareholders into account without requiring the knowledge of data on voting blocks. There is an intuitive meaning of power associated with our model: how important is a shareholder with respect to all other shareholders, or what is the relative voting power of a shareholder considering the dispersion of the rest of the votes? Applying the integrated model by virtue of Eq. (10) to H_{ij} yields $\tilde{H}_{ij}$. We are able to compute $\tilde{H}_{ij}$ for every shareholder in the sample without facing any computational restrictions (as opposed to the power indices). To summarize, the properties of our model make a sensible ranking of all shareholders according to their controlling power possible.

This concludes that our new measure of control merges crucial insights from the corporate finance literature and the game theoretic approach to voting while addressing their mentioned shortcomings. It should also be noted, that s_j represents the complementary of h_i: while the

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

latter represents the control seen from the point of view of the shareholders, the former reflects the control seen by the stocks.

4   Identifying the Backbone of Corporate Control

Based on the quantities introduced in the previous sections we are now in the position to proceed with the main aim of the paper, which is to investigate the concentration of control in the ownership networks at a global level. This means, qualitatively, that we have to identify those shareholders who can be considered to be in control of the market. In detail, we develop an algorithm that extracts the core subnetwork from the ownership network, which we call the *backbone*. This structure consists of the smallest set of the most powerful shareholders that, collectively, are potentially able to control a predefined fraction of the market in terms of value.

To this aim, in Sec. 4.1, we introduce a ranking of the shareholders based on the value of the portfolio they control, as measured by the integrated control value $\tilde{c}_i$, defined in Eq. (11). We are then able to compute how much value the top shareholders can potentially control, jointly, should they form a coalition. We call this notion *cumulative control*. Building on this knowledge, in Sec. 4.2, we extract the subnetwork of the most powerful shareholders and their (cumulatively) controlled stocks: the backbone. Sec. 4.3 presents a generalization of this backbone-extraction algorithm applicable to general weighted and oriented networks. The backbone structures of the analyzed countries are further investigated in Sec. 4.4. Different classification measures are introduced, allowing us to perform a cross-country analysis of how the control and value are globally distributed in the markets (Sec. 5.1) next to identifying who is holding the seat of power (Sec. 5.2).

4.1   Computing cumulative control

The first step of our methodology requires the construction of a Lorenz-like curve in order uncover the distribution of the value in a market. In economics, the Lorenz curve gives a graphical representation of the cumulative distribution function of a probability distribution. It is often used to represent income distributions, where the x-axis ranks the poorest $x\%$ of households and relates them to a percentage value of income on the y-axis.

Here, on the x-axis we rank the shareholders according to their importance and report the fraction they represent with respect to the whole set of shareholder. The y-axis shows the corresponding percentage of controlled market value. In detail, we relate the fraction of shareholders ranked by their integrated control value $\tilde{c}_i$, cf. Eqs. (3), (10) and (11), to the fraction of the total market value they collectively or cumulatively control.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

In order to motivate the notion of cumulative control, some preliminary remarks are required. Using the integrated control value to rank the shareholders means that we implicitly assume control based on the integrated fraction of control $\tilde{H}_{ij}$. This however is a potential value reflecting *possible* control. In order to identify the backbone, we take a very conservative approach to the question of what the *actual* control of a shareholder is. To this aim, we introduce a stringent threshold of 50%. Any shareholder with an ownership percentage $W_{ij} > 0.5$ controls by default. This strict notion of control for a single shareholder is then generalized to apply to the cumulative control a group of shareholders can exert. Namely by requiring the sum of ownership percentages multiple shareholders have in a common stock to exceed the threshold of cumulative control. Its value is equivalently chosen to be 50%.

We start the computation of cumulative control by identifying the shareholder having the highest $\tilde{c}_i$-value. From the portfolio of this holder, we extract the stocks that are owned at more than the said 50%. In the next step, the shareholder with the second highest $\tilde{c}_i$-value is selected. Next to the stocks individually held at more than 50% by this shareholder, additional stocks are considered, which are cumulatively owned by the top two shareholders at more than the said threshold value. See Fig. 7 for an illustrated example.

$U_{in}(n)$ is defined to be the set of indices of the stocks that are individually held above the threshold value by the n selected top shareholders. Equivalently, $U_{cu}(n)$ represents the set of indices of the cumulatively controlled companies. It holds that $U_{in}(n) \cap U_{cu}(n) = \emptyset$. At each step n, the total value of this newly constructed portfolio, $U_{in}(n) \cup U_{cu}(n)$, is computed:

$$v_{cu}(n) := \sum_{j \in U_{in}(n)} v_j + \sum_{j \in U_{cu}(n)} v_j. \tag{12}$$

Eq. (12) is in contrast to Eq. (5), where the total value of the stocks j is multiplied by the ownership percentage W_{ij}. The computation of cumulative control is described in steps $1 - 7$ (ignoring the termination condition in step 8) of Algorithm (1) on page 18. Consult the next section for more details.

Let n_{tot} be the total number of shareholders in a market and v_{tot} the total market value. We normalize with these values, defining:

$$\eta(n) := \frac{n}{n_{tot}}, \qquad \vartheta(n) := \frac{v_{cu}(n)}{v_{tot}}, \tag{13}$$

where $\eta, \vartheta \in (0, 1]$.

In Fig. (8) these values are plotted against each other for a selection of countries, yielding the cumulative control diagram, akin to a Lorenz curve (with reversed x-axis). As an example, a coordinate pair with value $(10^{-3}, 0.2)$ reveals that the top 0.1% of shareholders cumulatively control 20% of the total market value. The top right corner of the diagram represents 100%

J.B. Glattfelder, S. Battiston:
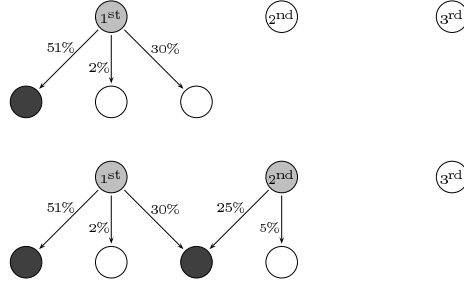The backbone of complex networks of corporations:
Who is controlling whom?

Figure 7: First steps in computing cumulative control: (*top panel*) selecting the most important shareholder (light shading) ranked according to the $\tilde{c}_i$-values and the portfolio of stocks owned at more than 50% (dark shading); in the second step (*bottom panel*), the next most important shareholder is added; although there are now no new stocks which are owned directly at more than 50%, cumulatively the two shareholder own an additional stock at 55%.
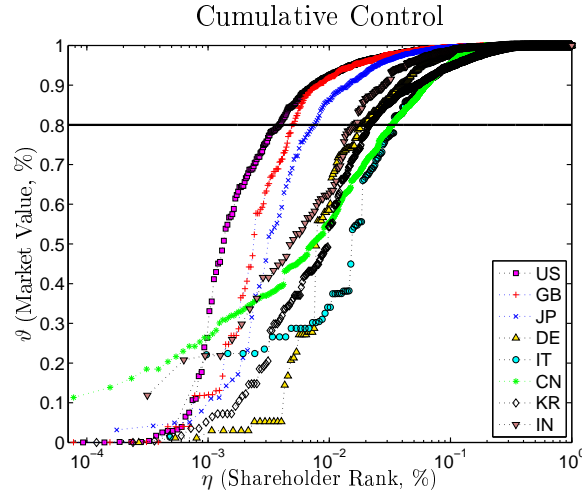

Figure 8: (Color online) Fraction of shareholders η, sorted by descending (integrated) control value $\tilde{c}_i$, cumulatively controlling ϑ percent of the total market value; the horizontal line denotes a market value of 80%; the diagram is in semi-log scale.

percent of the shareholders controlling 100% of the market value, and the first data point in the lower left-hand corner denotes the most important shareholder of each country. Different countries show a varying degree of concentration of control.

Recall that for every shareholder the ranking is based on all paths of control of any length along the direction of the arrows (indirect control). For every such reachable stock the importance of its direct co-shareholders is considered (against the direction of the arrows). Therefore our analysis is based on a genuine network approach which allows us to gain crucial information on every

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Algorithm 1 $\mathcal{BB}(\tilde{c}_1, \ldots, \tilde{c}_n, \delta, \hat{\vartheta})$

1: $\tilde{c} \leftarrow sort\_descending(\tilde{c}_1, \ldots, \tilde{c}_n)$
2: **repeat**
3:     $c \leftarrow get\_largest(\tilde{c})$
4:     $I \leftarrow I \cup index(c)$
5:     $PF \leftarrow stocks\_controlled\_by(I)$ (individually and cumulatively at more than δ)
6:     $PFV \leftarrow value\_of\_portfolio(PF)$
7:     $\tilde{c} \leftarrow \tilde{c} \setminus \{c\}$
8: **until** $PFV \geq \hat{\vartheta} \cdot total\_market\_value$
9: $prune\_network(I, PF)$

---

shareholder, which would otherwise be undetectable. In contrast, most other empirical studies start their analysis from a set of important stocks (e.g., ranked by market capitalization). The methods of accounting for indirect control (see Sec. 3.5) are, if at all, only employed to detect the so-called ultimate owners of the stocks. For instance, [24] studies the 10 largest corporations in 49 countries, [25] looks at the 20 largest public companies in 27 countries, [50] analyzes 2980 companies in nine East Asian countries, and [28] utilizes a set of 800 Belgian firms.

Finally, note that although the identity of the individual controlling shareholders is lost due to the introduction of cumulative control, the emphasis lies on the fact that the controlling shareholders are present in the set of the first n holders.

4.2 Extracting the backbone

Once the curve of the cumulative control is known for a market, one can set a threshold for the percentage of jointly controlled market value, $\hat{\vartheta}$. This results in the identification of the percentage $\hat{\eta}$ of shareholders that theoretically hold the power to control this value, if they were to coordinate their activities in corresponding voting blocks. As mentioned, the subnetwork of these power-holders and their portfolios is called the backbone. Here we choose the value $\hat{\vartheta} = 0.8$, revealing the power-holders able to control 80% of the total market value.

Algorithm (1) gives the complete recipe for computing the backbone. As inputs, the algorithm requires all the $\tilde{c}_i$-values, the threshold defining the level of (cumulative) control δ, and the threshold for the considered market value $\hat{\vartheta}$. As mentioned in the last section, steps $1 - 7$ are required for the cumulative control computation and δ is set to 0.5. Step 8 specifies the interruption requirement given by the controlled portfolio value being bigger than $\hat{\vartheta}$ times the total market value.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Finally, in step 9, the subnetwork of power-holders and their portfolios is pruned to eliminate weak links and further enhance the important structures: for each stock j, only as many shareholders are kept as the rounded value of s_j indicates, i.e., the (approximate) effective number of shareholders. E.g., if j has 5 holders but s_j is roughly three, only the three largest shareholders are considered for the backbone. In effect, the weakest links are removed.

4.3   Generalizing the method of backbone extraction

Notice that our method can be generalized to any directed and weighted network in which (1) a non-topological real value $v_j \geq 0$ can be assigned to the nodes (with the condition that $v_j > 0$ for at least all the leaf-nodes in the network) and (2) an edge from node i to j with weight W_{ij} implies that some of the value of j is transferred to i. In terms of physical systems, we do not seek a correspondence between the values v_j and the notion of a scalar potential. Instead, we think of the nodes as entities receiving material from the downstream nodes and transferring it to the upstream nodes without dissipation in proportion to the weights of the incoming links. Assume that the nodes which are associated with a value v_j produce v_j units of mass at time $t = 1$. Then the flow ϕ_i entering the node i from each node j at time t is the fraction W_{ij} of the mass produced directly by j plus the same fraction of the inflow of j:

$$\phi_i(t + 1) = \sum_j W_{ij} v_j + \sum_j W_{ij} \phi_i(t). \tag{14}$$

where $\sum_i W_{ij} = 1$ for the nodes that have predecessors and $\sum_i W_{ij} = 0$ for the root-nodes (sinks). In matrix notation, at the steady state, this yields

$$\phi = W(v + \phi). \tag{15}$$

The solution

$$\phi = (1 - W)^{-1} W v, \tag{16}$$

exists and is unique if $\lambda(W) < 1$. This condition is easily fulfilled in real networks as it requires that in each strongly connected component $\mathcal{S}$ there exists at least one node j such that $\sum_{i \in \mathcal{S}} W_{ij} < 1$. Or, equivalently, the mass circulating in $\mathcal{S}$ is also flowing to some node outside of $\mathcal{S}$. Notice that this does not imply that mass is lost in the transfer. Indeed, the mass is conserved at all nodes except at the sinks. Some of the nodes only produce mass (all the leaf-nodes but possibly also other nodes) at time $t = 1$ and are thus sources, while the root-nodes accumulate the mass. Note that it is straightforward to also define an equation for the evolution of the stock of mass present at each node.

The convention used in this paper implies that mass flows against the direction of the edges. This makes sense in the case of ownership, because although the cash allowing an equity stake in

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

a firm to be held flows in the direction of the edges, control (as defined by the integrated control value $\tilde{c}$) is transferred in the opposite direction, from the corporation to its shareholders. This is also true for the paid dividends. Observe that the integrated control value defined in Eq. (11) can be written in matrix notation as

$$\tilde{c} = \tilde{H}v = (1 - H)^{-1}Hv, \tag{17}$$

which is in fact equivalent to Eq. (16). This implies that for any node i the integrated control value $\tilde{c}_i = \sum_j \tilde{H}_{ij}v_j$ corresponds to the inflow ϕ_i of mass in the steady state.

Returning to the generic setting, let U_0 and E_0 be, respectively, the set of vertices and edges yielding the network. We define a subset $U \subseteq U_0$ of vertices on which we want to focus on (in the analysis presented earlier $U = U_0$). Let $E \subseteq E_0$ then be the set of edges among the vertices in U and introduce $\hat{\vartheta}$, a threshold for the fraction of aggregate flow through the nodes of the network. If the relative importance of neighboring nodes is crucial, H_{ij} is computed from W_{ij} by the virtue of Eq. (3). Note that H_{ij} can be replaced by any function of the weights W_{ij} that is suitable in the context of the network under examination. We now solve Eq. (10) to obtain the integrated value $\tilde{H}_{ij}$. This yields the quantitative relation of the indirect connections amongst the nodes. To be precise, it should be noted that in some networks the weight of an indirect connection is not correctly captured by the product of the weights along the path between the two nodes. In such cases one has to modify Eq. (8) accordingly.

The next step in the backbone extraction procedure is to identify the fraction of flow that is transfered by a subset of nodes. A systematic way of doing this was presented in Sec. 4.1 where we constructed the curve, (η, ϑ). A general recipe for such a construction is the following. On the x-axis all the nodes are ranked by their ϕ_i-value in descending order and the fraction they represent with respect to size of U is captured. The y-axis then shows the corresponding percentage of flow the nodes transfer. As an example, the first k (ranked) nodes represent the fraction $\eta(k) = k/|U|$ of all nodes that cumulatively transfer the amount $\vartheta(k) = (\sum_{i=1}^{k} \phi_i)/\phi_{tot}$ of the total flow. Furthermore, $\hat{\eta}$ corresponds to the percentage of top ranked nodes that pipe the predefined fraction $\hat{\vartheta}$ of all the mass flowing in the whole network. Note that the procedure described in Sec. 4.1 is somewhat different. There we considered the fraction of the total value given by the direct successors of the nodes with largest $\tilde{c}_i$. This makes sense due to the special nature of the ownership networks under investigation, where every non-firm shareholder (root-node) is directly linked to at least one corporation (leaf-node), and the corporations are connected amongst themselves.

Consider the union of the nodes identified by $\hat{\eta}$ and their direct and indirect successors, together with the links amongst them. This is a subnetwork $\mathcal{B} = (U^B, E^B)$, with $U^B \subset U$ and $E^B \subset E$ that comprises, by construction, the fraction $\hat{\vartheta}$ of the total flow. This is already a first possible definition of the backbone of (U, E). A discussion of the potential application of this procedure

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

to other domains, and a more detailed description of the generalized methodology (along with specific refinements pertaining to the context given by the networks) is left for future work. Viable candidates are the world trade web [8, 18, 51, 52], food-webs [4], transportation networks [53], and credit networks [54]

It should also be noted that in Sec. 4.1 we have introduced an additional threshold δ for the weights of the links which is needed in the context of corporate control. In the general case it can be set to zero. Returning to the specific context given by the data analyzed in this paper, one can vary the requirements that determine the backbone. For instance, one could focus on a predefined subset of listed companies, say the ten largest ones in the energy sector, and impose that the cumulative control over that set of stocks is $\hat{\vartheta} = 60\%$.

4.4   Defining classification measures

Markets are known to differ from one country to another in a variety of respects (see Sec. 1). They may however not look too different if one restricts the analysis to the distribution of local quantities, and in particular to the degree, as shown in Sec. 3.3. In contrast, at the level of the backbones, i.e., the structures where most of the value resides, they can look strikingly dissimilar, as seen for instance in the case of CN and JP, shown in Fig. 9. In order to attempt a classification of these diverse structures, we will make use of indicators built on the same quantities used to construct the backbone. Performing a cross-country analysis for these indicators gives new insights into the characteristics of the global markets.

In detail, the properties we are interested in and want to unveil are the concentration of control and value, next to the frequency of widely held companies. In the following, straightforward metrics reflecting these characteristics are defined. Let n_{st} and n_{sh} denote the number of stocks and shareholders in a backbone, respectively. As s_j measures the effective number of shareholders of a company, the average value

$$\overline{s} = \frac{\sum_{j=1}^{n_{st}} s_j}{n_{st}}, \tag{18}$$

is a good proxy characterizing the local patterns of ownership: the higher $\overline{s}$, the more dispersed the ownership is in the backbone, or the more common is the appearance of widely held firms. Furthermore, due to the construction of s_j, the metric $\overline{s}$ equivalently measures the local concentration of control.

In a similar vein, the average value

$$\overline{h} = \frac{\sum_{i=1}^{n_{sh}} h_i}{n_{sh}} = \frac{n_{st}}{n_{sh}}, \tag{19}$$

reflects the global distribution of control. A high value of $\overline{h}$ means that the considered backbone has very few shareholders compared to stocks, exposing a high degree of global concentration of

J.B. Glattfelder, S. Battiston:
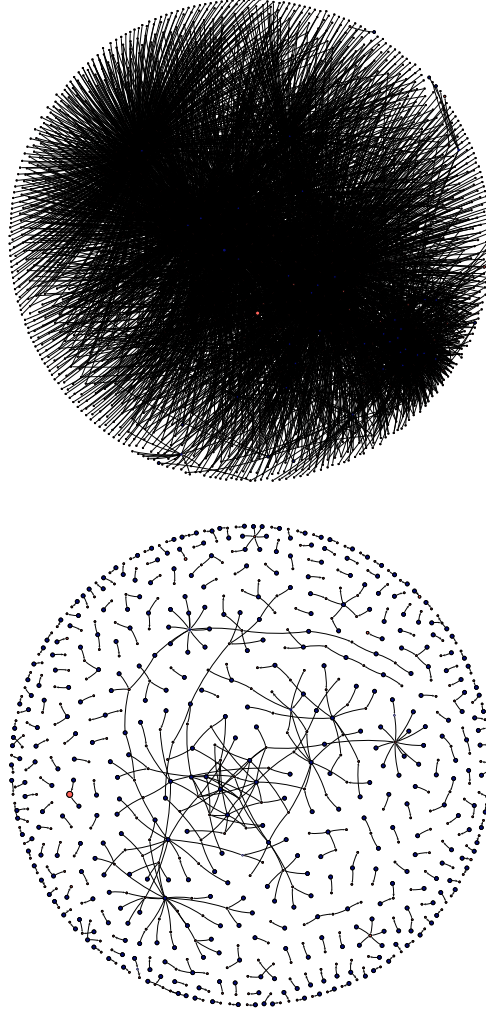The backbone of complex networks of corporations:
Who is controlling whom?

Figure 9: (*Top*) the backbone of JP; (*bottom*) the backbone of CN (for the complete set of backbone layouts consult [45]); the graph layouts are based on [55].

control. It is worth noting that the values n_{st} and n_{sh} are derived from the backbone and are hence network-related measures.

Recall that for the backbones to be constructed, a threshold for the controlled market value needed to be specified: $\hat{\vartheta} = 0.8$. In the cumulative control diagram seen in Fig. (8), this allows the identification of the number of shareholders being able to control this value. The value $\hat{\eta}$ reflects the percentage of power-holders corresponding to $\hat{\vartheta}$. To adjust for the variability introduced by the different numbers of shareholders present in the various national stock markets, we chose to normalize $\hat{\eta}$. Let n_{100} denote the smallest number of shareholders controlling 100% of the total

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Table 1: Classification measure values for a selection of countries; in Figs. 11 and 12 these values are plotted for all analyzed countries.

	η'	$\overline{s}$	$\overline{h}$
AU	0.82%	5.45	2.79
CA	3.32%	3.04	4.97
CH	5.97%	2.91	0.66
CN	9.21%	1.32	0.90
DE	3.22%	2.76	0.82
FR	3.96%	2.65	0.83
GB	0.89%	8.60	5.05
IN	5.27%	2.15	3.92
IT	6.10%	1.62	0.82
JP	1.93%	2.48	34.26
KR	2.25%	2.39	0.94
TW	5.00%	2.98	0.58
US	0.56%	8.56	15.39

market value v_{tot}, then

$$\eta' := \frac{\hat{\eta}}{n_{100}}. \tag{20}$$

A small value for η' means that there will be very few shareholders in the backbone compared to the number of shareholders present in the whole market, reflecting that the market value is extremely concentrated in the hands of a few shareholders. In essence, the metric η' is an emergent property of the backbone extraction algorithm and mirrors the global distribution of the value.

To summarize:

- $\overline{s}$ reflects local dispersion of control (at first-neighbor level, insensitive to value);

- $\overline{h}$ is an indicator of the global concentration of control (an integrated measure, i.e., derived by virtue of Eq. (10), at second-neighbor level, insensitive to value);

- η' is a global measure of the concentration of market value (an emergent quantity).

Table 1 shows the empirical values of these quantities for a selection of countries. In the following, the results of a cross-country analysis for the classification measures is given.

J.B. Glattfelder, S. Battiston:
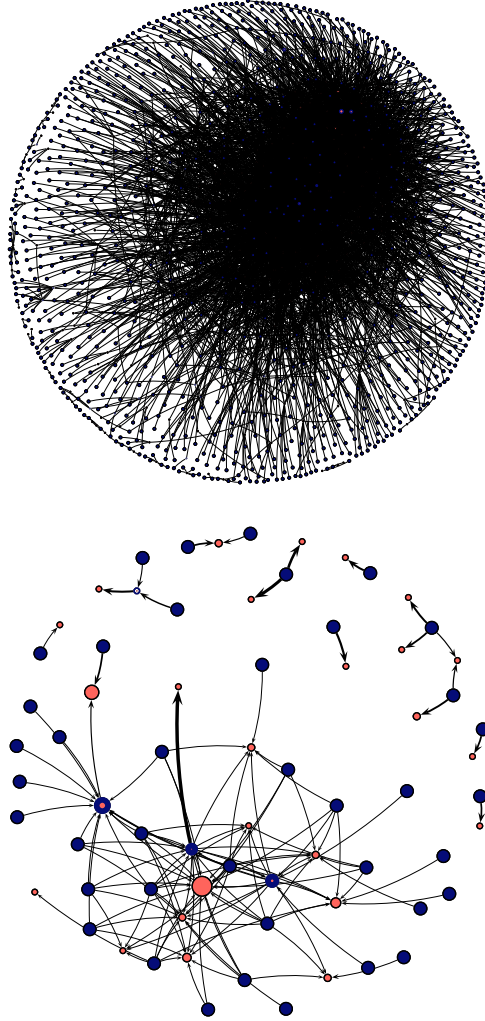The backbone of complex networks of corporations:
Who is controlling whom?

Figure 10: (*Top*) the ownership network of CH with 972 shareholders, 266 stocks and 4671 ownership relations; (*bottom*) the backbone of CH; firms are denoted by shaded nodes and sized by market capitalization, shareholders are black, whereas firms owning stocks themselves are represented by shaded nodes with thick bounding circles, arrows are weighted by the percentage of ownership value; the graph layouts are based on [55].

5   Analyzing the Backbones

In the last section, an algorithm for extracting the backbones of national markets, and measures reflecting their key characteristics, were given. But how relevant are these methods and how much of the properties of the real-world ownership networks they describe are captured?

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Fig. (10) shows the layout for the CH ownership network and the backbone, respectively. There is a big reduction in complexity by going to the backbone. Looking at the stocks left in the backbone, it is indeed the case that the important corporations reappear (recall that the algorithm selected the shareholders). We find a cluster of shareholdings linking, for instance, Nestlé, Novartis, Roche Holding, UBS, Credit Suisse Group, ABB, Swiss Re, Swatch. JPMorgan Chase & Co. features as most important controlling shareholder. The descendants of the founding families of Roche (Hoffmann and Oeri) are the highest ranked Swiss shareholders at position four. UBS follows as dominant Swiss shareholder at rank seven.

The backbone extraction algorithm is also a good test for the robustness of market patterns. The bow-tie structures (discussed in Sec. 3.1) in JP, KR, TW vanish or are negligibly small in their backbones, whereas in the backbones of the Anglo-Saxon countries (and as an outlier SE) one sizable bow-tie structure survives. This emphasizes the strength and hence the importance of these patterns in the markets of AU, CA, GB and the US.

But what about some of the findings in ownership patterns that have been previously reported in the literature? To see if we can recover some known observations, we analyze the empirical values for the "Widely Held" index defined in [25], where a value of one is assigned if there are no controlling shareholders, and zero if all firms in the sample are controlled. There is a threshold introduced, beyond which control is said to occur: the study is done with a 10% and 20% cutoff value. We find a 76.6% correlation (and a p-value for testing the hypothesis of no correlation of $3.2 \cdot 10^{-6}$) between $\overline{s}$ in the backbone and the 10% cutoff "Widely Held" index for the 27 countries it is reported for. The correlation of $\overline{s}$ in the countries' whole ownership networks is 60.0% ($9.3 \cdot 10^{-4}$). For the 20% cutoff, the correlation values are smaller. These relations should however be handled with care, as the study [25] is restricted to the 20 largest firms (in terms of market capitalization) in the analyzed countries and there is a twelve-year lag between the datasets in the two studies. Nevertheless, it is a reassuring sign to find such a high correlation with older proxies for the occurrence of widely held firms.

Having established that the backbones indeed successfully comprise important structures of the markets, and showing that one of the classification methods we propose confirms known results, we can proceed to investigate novel aspects of the ownership networks. As frequently mentioned in this paper, the lack of existing network-oriented analysis of the financial architecture of corporations in national markets leaves one question unaddressed: what is the *global* concentration of control?

5.1   Global concentration of control

We utilize the measures defined in Eqs. (18), (19) and (20), to classify the 48 backbones. To recapitulate, $\overline{s}$ is a local measure for the dispersion of control. A large value indicates a high
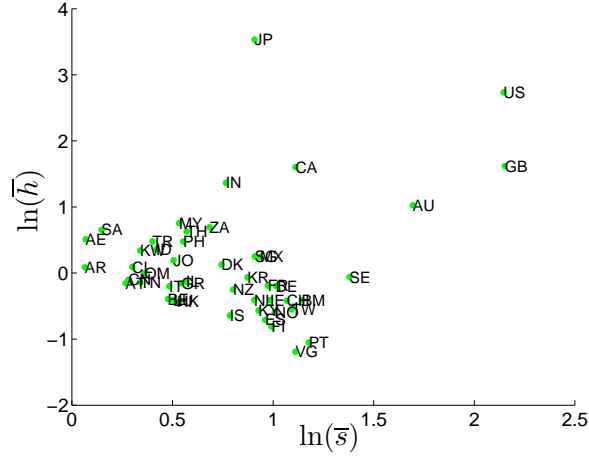
J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

Figure 11: Map of control: local dispersion of control, $\overline{s}$, is plotted against global concentration of control, $\overline{h}$, for 48 countries.

presence of widely held firms. $\overline{h}$ is a second-neighbor quantity sensitive to the concentration of global control. Large values are indicative that the control of many stocks resides in the hands of very few shareholders. Finally, η' is a global variable related to the (normalized) percentage of shareholders in the backbone. It hence measures the concentration of value in a market, as a low number means that very few shareholders are able to control 80% of the market value.

In Fig. 11 the log-values of $\overline{s}$ and $\overline{h}$ are plotted against each other. The $\overline{s}$-coordinates of the countries are as expected [25]: to the right we see the presence of widely held firms (i.e., the
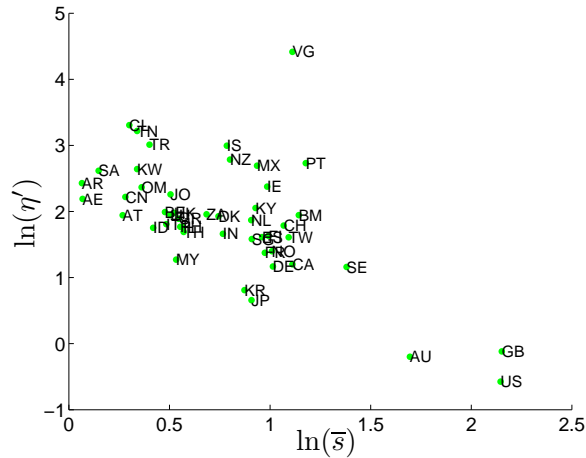

Figure 12: Map of market value: local dispersion of control, $\overline{s}$, is plotted against global concentration of market value, η', for 48 countries.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

local dispersion of control) for the Anglo-Saxon countries AU, GB and the US. FR, IT, JP are located to the left, reflecting more concentrated local control. However, what is astonishing is that there is a counterintuitive trend to be observed in the data: the more local control is dispersed, the higher the global concentration of control becomes. In essence, what looks like a democratic distribution of control from close up, by taking a step back, actually turns out to warp into highly concentrated control in the hands of very few shareholders. On the other hand, the local concentration of control is in fact widely distributed amongst many controlling shareholder. Comparing with Fig. 2, where idealized network configurations are illustrated, we conclude that the empirical patterns of local and global control range from the type (B) to type (D), with JP combining local and global concentration of control. Interestingly, type (A) and (C) constellations are not observed in the data.

In Fig. 12 the log-values of $\overline{s}$ and η' are depicted. What we concluded in the last paragraph for control is also true for the market value: the more the control is locally dispersed, the higher the concentration of value that lies in the hands of very few controlling shareholders, and vice versa.

We can also compare the $\overline{s}$ and $\overline{h}$ values measured for the backbones with the corresponding values of the total ownership networks, $\overline{s}_{tot}$ and $\overline{h}_{tot}$. We find that

$$\overline{s} < \overline{s}_{tot}. \tag{21}$$

This fact, that the widely held firms are less often present in the national backbones, means that the important shareholders (able to control 80% of the market value) only infrequently invest in corporations with dispersed ownership. Note that the pruning scheme used in the construction of the backbone (introduced at the end of Sec. 4.2) approximates s_j to the nearest integer. This can reduce the value of $\overline{s}$ in the backbone maximally by 0.5. In contrast, in our data (with the exception of ES) the relation $\overline{s}_{tot} - \overline{s} \gg 0.5$ holds, indicating that there is indeed a tendency of power-holders to avoid widely held firms, accounting for their less frequent appearance in the backbones.

We also find that

$$\overline{h} > \overline{h}_{tot}. \tag{22}$$

This means that there is a higher level of global control in the backbone, again implying that widely held firms occur less often in the backbone. In addition, looking at the ranges of $\overline{h}_{tot} \in [0.06, 1.09]$ and $\overline{h} \in [0.3, 34.26]$, reveals a higher cross-country variability in the backbone. In essence, the algorithm for extracting the backbone in fact amplifies subtle effects and unveils key structures.

We realize that the two figures discussed in this section open many questions. Why are there outliers to be observed: JP in Fig. 11 and VG in Fig. 12? What does it mean to group countries

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

according to their $\overline{s}$, $\overline{h}$ and η' coordinates and what does proximity imply? What are the implications for the individual countries? We hope to address such and similar questions in future work.

5.2 The seat of power

Having identified the important shareholders in the global markets, it is now also possible to address the following questions. Who holds the power in an increasingly globalized world? How important are individual people compared to the sphere of influence of multinational corporations? How eminent is the influence of the financial sector? By look in detail at the identity of the power-holders featured in the backbones, we address these issues next.

If one focusses on how often the same power-holders appear in the backbones of the 48 countries analyzed, it is then possible to identify the global power-holders. Unsurprisingly, they turn out to be mostly multinational corporations in the banking and insurance sectors. Below is a top-ten list, comprised of the companies' name, activity, country the headquarter is based in, and ranked according to the number of times it is present in different countries' backbones.

1. The Capital Group Companies; investment management; US; 36;

2. Fidelity Management & Research; investment products and services; US; 32;

3. Barclays PLC; financial services provider; GB; 26;

4. Franklin Resources; investment management; US; 25;

5. AXA; insurance company; FR; 22;

6. JPMorgan Chase & Co.; financial services provider; US; 19;

7. Dimensional Fund Advisors; investment management; US; 15;

8. Merrill Lynch & Co.; investment management; US; 14;

9. Wellington Management Company; investment management; US; 14;

10. UBS; financial services provider; CH; 12.

As mentioned, the prevalence of companies in the financial and insurance sectors is perhaps not very surprising. After all, Capital Group Companies is one of the world's largest investment management organizations with assets under management in excess of one trillion USD. However, it is an interesting observation that all the above mentioned corporations appear as prominent controlling shareholders in the various backbones they are present in.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

The dominance of US American companies seems slowly to be contested: next to Barclays PLC (GB), AXA (FR) and UBS (CH), we find Deutsche Bank (DE), Brandes Investment Partners (CA), Société Générale (FR), Credit Suisse Group (CH), Schroders PLC (GB), Allianz (DE) in the top 21 positions. The government of Singapore is at rank 25. HSBC Holdings PLC (HK/GB), the world's largest banking group, only appears at position 26.

In addition, large multinational corporations outside of the finance and insurance industry do not act as prominent shareholders and only appear in their own national countries' backbones as controlled stocks. For instance, Exxon Mobil, Daimler Chrysler, Ford Motor Company, Siemens, Unilever.

The observation that individual people do not appear as multinational power-holders is perhaps also not surprising. Indeed, most countries' backbones do not have people appearing in the top-ten list of shareholders. In the US backbone, we find one person ranked at ninth position: Warren E. Buffet. William Henry Gates III is next, at rank 26. In DE the family Porsche/Piech and in FR the family Bettencourt are power-holders in the top ten. For the tax-haven KY one finds Kao H. Min (who is placed at number 140 in the Forbes 400 list) in the top ranks.

6   Summary and Conclusion

We have developed a methodology to extract the backbone of corporate control networks, that is a subnetwork where most of the control and the economic value resides. In this procedure the indirect control along all ownership pathways is fully accounted for. The methodology applies in general to networks with weighted and directed links in which nodes are associated with a scalar quantity.

We can interpret such networks as systems in which mass is created at some nodes and transferred to the nodes upstream. The amount of mass flowing along a link from node i to node j is given by the scalar quantity associated with the node j, times the weight of the link, $W_{ij} v_j$. The backbone corresponds to the subnetwork in which a preassigned fraction of the total flow of the system is transfered.

From a network theoretic point of view, we extended the notions of degree to more suitable measures that take into account the relative weight of the links with respect to the links of second-order neighbors. Nodes were associated with non-topological state variables given by the market capitalization size of the firms. We ranked the shareholders according to the value they can control and we constructed the subset of shareholders which collectively control a given fraction of the economic value in the market. We further introduced some measures aimed at classifying the backbone of the different markets in terms of local and global concentration of control and value.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

With respect to the literature addressing the empirical analysis of economic networks, this paper presents the first extensive cross-country investigation of the control of corporations based on ownership relations and market capitalization values in 48 national stock markets.

We find that each level of detail (i.e., topology, weights and direction, value of nodes) in the analysis uncovers new features in the ownership networks. Incorporating the direction of links in the study reveals bow-tie structures in the network. Including value allows us to identify who is holding the power in the global stock markets.

With respect to other studies in the economics literature, next to proposing a new model for estimating control from ownership, we are able to recover previously observed patterns in the data, namely the frequency of widely held firms in the various countries studied. Indeed, it has been known for over 75 years that the Anglo-Saxon countries have the highest occurrence of widely held firms [56]. This statement, that the control of corporations is dispersed amongst many shareholders, invokes the intuition that there exists a multitude of owners that only hold a small amount of shares in a few companies. However, in contrast to such intuition, our main finding is that a local dispersion of control is associated with a global concentration of control and value. This means that only a small elite of shareholders continually reappears as the controlling entity of all the stocks, without ever having been previously detected or reported on. On the other hand, in countries with local concentration of control (mostly observed in European states), the shareholders tend to only exert control over a single corporation, resulting in the dispersion of global control and value.

Finally, we also observe that the US financial sector holds the seat of power at an international level. It will remain to be seen, if the continued unfolding of the current financial crises will tip this balance of power, as the US financial landscape faces a fundamental transformation in its wake.

7   Acknowledgements

References

[1] R. Albert, H. Jeong, and A. Barabási, Nature **401**, 130 (1999).

[2] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

[3] M. Newman, D. Watts, and S. Strogatz, Proc. Natl. Acad. Sci. **99**, 2566 (2002).

[4] D. Garlaschelli, G. Caldarelli, and L. Pietronero, Nature **423**, 165 (2003).

[5] S. Dorogovtsev and J. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (2003).

[6] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, 2007).

[7] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci. **101**, 3747 (2004).

[8] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).

[9] J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski, Phys. Rev. E **71**, 065103.1 (2005).

[10] D. Garlaschelli, S. Battiston, M. Castri, V. D. P. Servedio, and G. Caldarelli, Physica A **350**, 491 (2005).

[11] G. De Masi, G. Iori, and G. Caldarelli, Phys. Rev. E **74**, 66112 (2006).

[12] R. Albert and A. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[13] M. Newman, S. Strogatz, and D. Watts, Phys. Rev. E **64**, 26118 (2001).

[14] S. Battiston and M. Catanzaro, Eur. Phys. J. B **38**, 345 (2004).

[15] S. Battiston, J. F. Rodrigues, and H. Zeytinoglu, Adv. Complex Syst. **10**, 29 (2005).

[16] G. Bonanno, G. Caldarelli, F. Lillo, and R. Mantegna, Phys. Rev. E **68**, 46130 (2003).

[17] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, and A. Kanto, Phys. Rev. E **68**, 56110 (2003).

[18] M. Serrano and M. Boguñá, Phys. Rev. E **68**, 15101 (2003).

[19] D. Garlaschelli and M. Loffredo, Physica A **355**, 138 (2004).

[20] P. Windolf, *Corporate Networks in Europe and the United States* (Oxford University Press, 2002).

[21] J. E. Stiglitz, J. of Money, Credit and Banking **17**, 133 (1985).

[22] K. M. Eisenhardt, Acad. Manage. Rev. **14**, 57 (1989).

[23] A. Shleifer and R. W. Vishny, J. Finance **52**, 737 (1998).

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

[24] R. La Porta, F. L. de Silanes, A. Shleifer, and R. Vishny, J. Polit. Econ. **106**, 1113 (1998).

[25] R. La Porta, F. L. de Silanes, and A. Shleifer, J. Finance **54** (1999).

[26] F. Brioschi, L. Buzzacchi, and M. Colombo, J. Bank. Financ. **13**, 747 (1989).

[27] D. Flath, Econ. Lett. **38**, 223 (1992).

[28] A. Chapelle, Corporate Ownership and Control **15** (2005).

[29] A. Chapelle and A. Szafarz, Physica A **355**, 509 (2005).

[30] H. Almeida and D. Wolfenzon, J. Finance **61**, 2637 (2006).

[31] M. Granovetter, *Ind. Corp. Change* (Oxford University Press, 1995), chap. Coase Revisited: Business Groups in the Modern Economy.

[32] L. S. Shapley and M. Shubik, Am. Polit. Sci. Rev. **48**, 787 (1954).

[33] J. F. Banzhaf, Rutgers Law Rev. **19**, 317 (1965).

[34] D. Leech, Management Sci. **34**, 509 (1988).

[35] B. Kogut and G. Walker, Am. Sociol. Rev. **66**, 317 (2001).

[36] R. Corrado and M. Zollo, Ind. Corp. Change **15**, 319 (2006).

[37] G. Fagiolo, Phys. Rev. E **76**, 26107 (2007).

[38] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, S. Stata, A. Tomkins, and J. Wiener, Computer Networks **33**, 309 (2000).

[39] S. Vitali, J. B. Glattfelder, and S. Battiston (2008), working paper.

[40] R. C. Feenstra, T.-H. Yang, and G. G. Hamilton, J. Int. Econ. **48**, 71 (1999).

[41] R. Dore, New Polit. Economy **7**, 115 (2002).

[42] O. Herfindahl, *Copper Costs and Prices: 1870 – 1957* (John Hopkins University Press, Baltimore, 1959).

[43] J. Cubbin and D. Leech, Econ. J. **93**, 351 (1983).

[44] B. Derrida and H. Flyvbjerg, J. Phys. **A**, 1003 (1986).

[45] Supplementary material available at `http:// www.sg.ethz.ch/people/jglattfelder/online`.

J.B. Glattfelder, S. Battiston:
The backbone of complex networks of corporations:
Who is controlling whom?

[46] The Deminor Group, Tech. Rep., A report prepared for the Association of British Insurers (2005), URL http://www.abi.org.uk/Bookshop/.

[47] D. Leech, The Warwick Economics Research Paper Series 644, University of Warwick, Department of Economics (2002), URL http://ideas.repec.org/p/wrk/warwec/644.html.

[48] S. Prigge, SSRN eLibrary (2007), URL http://ssrn.com/paper=1100619.

[49] D. Leech and J. Leahy, Econ. J. **101**, 1418 (1991).

[50] S. Claessens and S. Djankov, J. Financ. Econ. **58**, 81 (2000).

[51] J. Reichardt and D. White, Eur. Phys. J. B **60**, 217 (2007).

[52] G. Fagiolo, J. Reyes, and S. Schiavo, Physica A **387** (2008).

[53] C. Kühnert, D. Helbing, and G. West, Physica A **363**, 96 (2006).

[54] S. Battiston, D. Delli Gatti, M. Gallegati, B. Greenwald, and J. Stiglitz, J. Econ. Dyn. Control **31**, 2061 (2007).

[55] M. M. Geipel, Int. J. Mod. Phys. C **18**, 1537 (2007).

[56] A. Berle and G. Means, *The Modern Corporation and Private Property* (Mac Millan, N.Y., 1932).