Control and voting power in corporate networks: Concepts and computational aspects

Yves Crama¹ and Luc Leruth²

March 2005

¹HEC Management School, University of Liège, Boulevard du Rectorat 7 (B31), 4000 Liège, Belgium, Y.Crama@ulg.ac.be
²Fiscal Affairs Department, International Monetary Fund, 700 19th Street, Washington D.C. 20431, U.S.A., LLeruth@imf.org
Abstract

This paper proposes to rely on power indices to measure the amount of control held by individual shareholders in corporate networks. The value of the indices is determined by a complex voting game viewed as the composition of interlocked weighted majority games; the compound game reflects the structure of shareholdings. The paper describes an integrated algorithmic approach which allows to deal efficiently with the complexity of computing power indices in shareholding networks, irrespective of their size or structure. In particular, the approach explicitly accounts for the presence of float and of cyclic shareholding relationships. It has been successfully applied to the analysis of real-world financial networks.

Keywords: finance, corporate networks, game theory, power indices, simulation

Disclaimer: The views expressed in this paper do not necessarily reflect those of the IMF or of its Board of Directors.
1 Introduction

The structure of corporate networks has recently received growing attention, in particular in relation to the principal-agent issues arising between shareholders and managers. The analysis of these issues is complicated by the existence of two distinct systems of governance in financial markets: the outsider system, characterized by simple structures with a single layer of shareholders (usually individuals, but also investment funds) and a market for control; and the insider system, characterized by complex, multilayered structures, where firms hold significant shares of other firms and where the market for control is not fully transparent. According to this typology, the world can broadly be divided into separate groups including the USA and the United Kingdom on the one hand (the outsider system), and most of the rest of the world on the other hand (the insider system). While the distinction between the two systems can at times appear arbitrary, it has been widely used in the literature.

If the features of the insider system make it especially difficult to analyze shareholding relationships, the notion of control is often not precisely defined even in the simpler case of the outsider system (see La Porta, Lopez-De-Silanes, and Shleifer (1999) or Crama, Leruth, Renneboog, and Urbain (2003)). Yet, a number of papers address this issue and focus on the need to distinguish between “control” and “ownership”. Some of these papers are mostly empirical and descriptive, while others attempt to provide a game theoretic framework for their analysis. In order to explain more precisely the focus of our contribution, it is necessary to start with a brief review of these previous approaches (more complete overviews of the literature are presented in Shleifer and Vishny (1997) and Crama, Leruth, Renneboog, and Urbain (2003)).

La Porta, Lopez-De-Silanes, and Shleifer (1999) analyze ownership concentration in 27 wealthy economies. They focus on the issue of “controlling ownership” and find that very different patterns of ownership concentration occur in different countries. This confirms previous observations that “widely held” companies (i.e. companies with very dispersed ownership, as they are found in the outsider systems of governance mentioned above) do not constitute the only prevailing paradigm. The definition of a controlling shareholder in La Porta, Lopez-De-Silanes, and Shleifer (1999) relies on the concept of a cutoff point, i.e. a critical percentage of shares beyond which the largest shareholder is deemed to possess effective control over the firm’s decisions (although the paper also explores the impact of alternative indicators of control). Their study mostly uses as cutoff a shareholding exceeding 20% of total shares. The concept of cutoff point has been adopted by several authors (going back to Berle and Means (1932)), though the actual percentage used in different studies may vary from 10% to 30%.

Claessens, Djankov, and Lang (2000) and Claessens, Djankov, Fan, and Lang (2002) analyze the complex pyramidal structures found in Asian markets and observe that control is achieved through pyramidal structures in many countries. In both papers, the authors rely on a definition of control that was once suggested by the classical analysis of Berle and Means (1932).
which only takes voting rights into account (while their definition of ownership is based on cash-flow rights). In short, they consider that, along a chain of voting rights from a parent corporation $a$ to firm $b$, to firm $c$, ..., to firm $t$, the weakest link (i.e., smallest shareholding) in the chain defines the extent of the control of $a$ on $t$. Control rights along different chains are assumed to be purely additive. This approach can be illustrated by the following example (from Claessens, Djankov, and Lang (2000) p. 91).

**Example 1.** Suppose that firm $a$ holds 11% of the stock of firm $b$ and 25% of firm $c$, as represented in Figure 1. Moreover, firm $b$ owns 21% of firm $t$ and firm $c$ owns 7% of firm $t$. Then, according to the previous definition, firm $a$ controls $11\% + 7\% = 18\%$ of firm $t$. □

![Figure 1: The network for Example 1](image)

Thus, in the above studies, “control” is assimilated to “shareholding”. Irrespective of how this assimilation is done, it necessarily leads to a crude measure of control which has been challenged by several authors. To understand why, consider again Example 1 and assume that the remaining shares of $t$ (72%) are concentrated in the hands of another firm $d$. It would then be difficult to maintain the claim that $a$ controls 18% of $t$ since, for all practical purposes, $d$ actually has complete control over $t$.

Although our counter-example is quite simplistic (and could easily be ruled out by appropriate ad hoc mechanisms), it provides useful insights into some of the key features that should be expected from any realistic measure of control: first, such measure must take the whole distribution of shares into account, since the amount of control owned by a shareholder typically depends on the holdings of the other shareholders; moreover, a sound measure of control will probably not grow in direct proportion to shareholdings, and, in any case, the measure will undergo a sharp discontinuity when a majority of the shares are transferred to a unique owner.

In view of these considerations, it is not surprising that several authors have proposed to use game-theoretical power indices in order to measure control in corporate groups. This suggestion

In these approaches, shareholders are modeled as players in a simple (or voting) game (see e.g. Bilbao (2000), Felsenthal and Machover (1998), Owen (1995), Shubik (1982)). Recall that a simple game is a monotonically increasing function \( v : 2^N \rightarrow \{0,1\} \), where \( N = \{1,2,\ldots,n\} \) is a finite set of players and \( 2^N \) is the power set of \( N \). Equivalently, a simple game can be modeled as a Boolean function \( f_v : \{0,1\}^n \rightarrow \{0,1\} \): here, the variables of \( f_v \) are in 1-to-1 correspondence with the players of the game (variable \( x_i \) takes value 1 exactly when player \( i \) votes “Yes”) and the value of the function reflects the outcome of the vote for each vector of individual votes. More precisely, for all \( X = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n \),

\[
\begin{align*}
    f_v(X) &= 1 \quad \text{if } v(\{i : x_i = 1\}) = 1, \\
    &= 0 \quad \text{otherwise},
\end{align*}
\]

where \( \{i : x_i = 1\} \) is the index set of the players who vote “Yes”. This equivalent model allows, in particular, to speak of the composition of \( k \) simple games \( f_1, f_2, \ldots, f_k \) by another game \( f \): this is just the game \( f(f_1, f_2, \ldots, f_k) \). For simplicity, we shall often use the same notation for the game \( v \) viewed either as a set-function on \( 2^N \) or as a Boolean function on \( \{0,1\}^n \).

A main theme of study in game theory is the computation of the amount of power held by the players, and several indices have been proposed in the literature for this purpose (we refer to Bilbao (2000), Dubey and Shapley (1979), Felsenthal and Machover (1998), Owen (1995) or Shubik (1982) for detailed presentations). In the context of corporate networks, power indices are expected to reflect the relative capacity of each shareholder (be it a firm or an individual) to impose its will to a target company, measured either by its ability to form coalitions with other shareholders to win a vote, or by its ability to change the final outcome of a vote by swinging its own vote. Lengthy discussions regarding the relevance of power indices for the analysis of control in corporate structures and various applications can be found in the references cited above.

When applied to corporate networks, a main limitation of previous studies on power indices is that they remain confined to the analysis of the weighted majority game defined by a single layer of shareowners (i.e. they only deal with the case of an outsider system of governance, as described above). In this simplest case, a target firm \( t \) has \( n \) direct shareholders \( 1, 2, \ldots, n \), and firm \( i \) detains the fraction \( w_i \) of the shares of \( t \) \((i = 1, \ldots, n)\). When the shareholders cast their votes \((x_1, x_2, \ldots, x_n) \in \{0,1\}^n \), the outcome of the procedure is determined by computing the sum \( \sum_{i=1}^n w_i x_i \): if this sum is larger than 0.5, then the outcome is 1; otherwise, it is 0 (quotas other than 0.5 would be similarly handled).

This limitation is in sharp contrast with the stream of research on “control” vs. “ownership”
represented by the papers of Claessens et al. (2000, 2002) or La Porta et al. (1999). Indeed, these and related papers clearly identify and systematically investigate (albeit without reference to formal game-theoretic models) pyramidal, multilayered ownership structures, such as they can be found in the insider system of governance.

The purpose of the present paper is to describe an algorithmic approach for the estimation of power indices in corporate networks, whether simple and monolayered, or complex and interlocking. We focus on the computation of the Banzhaf index (see Banzhaf (1965)), which we define more precisely in Section 2 after a presentation of our game-theoretic model. In Section 3, we propose a methodology to address a number of difficult issues which arise when dealing with pyramidal ownership structures: algorithmic exploitation of the underlying graph model, identification of “ultimate shareholders”, models for the float of small shareholders, cyclic cross-holdings, etc. We demonstrate in Section 4 the practical feasibility of the approach for handling large, real-world databases, compiled from either the insider or the outsider system of governance. Finally, in Section 5, we provide a number of applications of our approach to the analysis of various financial issues.

2 Models

2.1 Graphs and games

In order to describe our formal model of corporate networks, we need some graph-theoretic concepts and terminology. We mostly follow the terminology of Bang-Jensen and Gutin (2000). We suppose that the shareholding relationships among a set of firms $V$ are described by the network $G = (V, A; w)$, where $(V, A)$ is a directed graph with vertex set $V$ and with arc set $A \subseteq V \times V$, and where $w$ is a real-valued positive mapping on the arc-set: for each arc $(i, j) \in A$, $w_{ij} > 0$ represents the fraction of the stock of firm $j$ owned (directly) by firm $i$. When $(i, j) \in A$, we say that $i$ is a predecessor (or direct shareholder) of $j$ and that $j$ is a successor of $i$. A walk is a sequence of vertices $(i_1, i_2, \ldots, i_k)$ such that $(i_r, i_{r+1}) \in A$ for $r = 1, \ldots, k - 1$. If all vertices of a walk are distinct, then we say it is a path; a cycle is a path such that $i_1 = i_k$. Vertex $i$ is an ancestor of vertex $j$ if there is a path from $i$ to $j$ in $G$. A strong component of $G$ is a maximal subset $S$ of vertices such that, for every pair $i, j$ of distinct vertices in $S$, $i$ is an ancestor of $j$ and $j$ is an ancestor of $i$.

For all $j \in V$, we denote by $V_j$ the set of predecessors of firm $j$: $V_j = \{ i : (i, j) \in A \}$. When $V_j = \emptyset$, we say that $j$ is a source of the network; we denote by $N = \{ 1, 2, \ldots, n \}$ the set of sources of $G$. Conversely, a target (or sink) of the network is a firm which does not have any successor in $G$.

The direct game $g_j$ associated with each firm $j \in V \setminus N$ is the weighted majority game with player-
set $V_j$ and with weights $\{w_{ij} : (i, j) \in A\}$. Thus, the direct game $g_j$ provides a simple model for the decision-making process at a shareholder meeting of firm $j$.

In order to extend the framework to more complex structures, let us first consider the special case where $G$ is acyclic, i.e. where $G$ does not contain any cycles. In this case, the network necessarily contains one or several sources, i.e. $N \neq \emptyset$. Then, an indirect game $v_j$ can be associated with each firm $j \in V$ by composition of the direct games associated with its predecessors. More precisely, we define the player-set of $v_j$ to be $N = \{1, 2, \ldots, n\}$ and we associate a $0 - 1$ variable $x_j$ with each $j \in N$. If $j$ is a source, then $v_j$ is defined as $v_j(X) = x_j$ for all $X = (x_1, x_2, \ldots, x_n) \in \{0, 1\}^n$. Otherwise, if $i_1, i_2, \ldots, i_k$ denote the direct shareholders of $j$, then $v_j$ is recursively defined as

$$v_j(X) = g_j(v_{i_1}(X), v_{i_2}(X), \ldots, v_{i_k}(X)) \quad \text{for all } X \in \{0, 1\}^n. \quad (1)$$

The computation of the indirect games requires complete knowledge of the shareholding structure in the corporate network, i.e. the following condition must hold:

$$\sum_{i \in V_j} w_{ij} = 1 \quad \text{for all } j \in V \setminus N \quad (2)$$

(if firm $j$ is not a source, then all its shareholders must appear in the network). When condition (2) holds, we say that $G$ and all associated indirect games are complete. If the condition fails for a firm $j \in V \setminus N$, then $G$ and all games which depend on $j$ are called incomplete.

The quantity

$$fl_j \overset{\text{def}}{=} 1 - \sum_{i \in V_j} w_{ij} \quad (3)$$

defines the float of firm $j \in V \setminus N$ (sometimes, we also call “float” the set of unidentified shareholders of the firm). The float can be interpreted as missing data: when $fl_j$ is nonzero, we do not know how to compute the direct game associated with $j$ but the definition of this game remains conceptually valid (some of the shareholdings are just unknown). We shall explain in Section 3.2 how we deal with the float in our computations.

The most complex situation arises when the network contains cycles: in this case, Eq. (1) cannot be used to define the indirect games (the recursion becomes cyclic) and the intuition underlying the game-theoretic model is in jeopardy. Since cyclic corporate structures do occur in practice, we shall propose various ways to deal with cyclic networks in Section 3.4.

\[\text{For the interested reader, this is identical to the construction of a feed-forward neural network of threshold units, as defined for instance in Anthony (2001).}\]
2.2 Banzhaf power index

If \( v \) is a simple game on the player-set \( N = \{1, 2, \ldots, n\} \), the (non-normalized) Banzhaf index of player \( j \in N \) is the quantity

\[
Z_v(j) = \frac{1}{2^{n-1}} \cdot \sum_{T \subseteq N:j \in T} (v(T) - v(T \setminus \{j\})))
\]

(4)

\[
= \frac{1}{2^{n-1}} \cdot \sum_{X \in \{0,1\}^n : x_j = 1} (v(X) - v(X - e_j)),
\]

(5)

where \( e_j \) is the \( j \)-th \( n \)-dimensional unit vector.

When \( G = (V, A; w) \) is an acyclic complete network, \( t \) is an arbitrary vertex and \( j \) is a source of \( G \), we denote by \( Z_G(j, t) \) the Banzhaf index of firm \( j \) in the indirect game \( v_t \) associated with \( t \). We drop the subscript \( G \) and we simply write \( Z(j, t) \) whenever this does not lead to confusion.

The Banzhaf index of player \( j \) can be usefully interpreted as the probability that \( j \) can change the outcome of a vote by changing his own vote from 1 to 0, assuming that all players vote randomly and are equally likely to vote either 0 or 1 (cf. Banzhaf (1965) and the pioneering work by Penrose (1946); see also Dubey and Shapley (1979), Owen (1995), etc.). We use the Banzhaf index as a measure of the amount of \textit{a priori} voting power, or control, held by a firm. As mentioned earlier, this proposal has been formulated and exploited by numerous researchers, and we do not intend to justify it further in this paper. Let us simply point out that recent work by Leech (2002a) tends to demonstrate that the Banzhaf index may be better suited than the Shapley-Shubik index in the corporate context.

3 Main algorithmic issues

As mentioned before, all previous computational studies on power indices have dealt with very simple structures, usually restricted to majority games played by one layer of shareholders, possibly under special majority rules, or to two-level compound games; see e.g. Laruelle and Widgrén (1998), Leech (2002b), Owen (1975), Shapley and Shubik (1954), Strand (2003), etc.

The Banzhaf indices of a weighted majority game on \( n \) players can be computed by dynamic programming or generating function techniques (see e.g. Brams and Affuso (1976), Bilbao (2000)). The complexity of these algorithms, however, is not polynomial but only \textit{pseudo-polynomial}. In fact, computing the exact value of the Banzhaf indices of a weighted majority game is a hard problem; more precisely, it belongs to the class of \#P-hard problems (Garey and Johnson (1979)). As a consequence, exact algorithms for this problem become prohibitively slow when the number of players grows large, and one must resort in this case to approximation techniques (see e.g. Leech (2003)) or to Monte Carlo simulation (as originally suggested by Mann and Shapley (1960)).

Owen (1978, 1995) considered the Banzhaf index of compound games. In our terminology, his
results apply to an acyclic network $G$ with a single target $t$, and where each firm $i \neq t$ has exactly one successor $s(i)$. Owen (1978) observed that, in the game $v_t$ associated with $G$, the Banzhaf index of a source-player $i$ can be recursively computed from the relation

$$Z_G(i, t) = Z_G(i, s(i)) \cdot Z_H(s(i), t), \quad (6)$$

where $Z_G(i, s(i))$ is the Banzhaf index of $i$ in the direct game $g_{s(i)}$, and $H$ is the network obtained by deleting all ancestors of $s(i)$ from $G$; see also Dubey, Einy and Haimanko (2003). This allows to reduce the computation of $Z_G$ to the computation of Banzhaf indices of weighted majority games. Note however that the simple relation (6) does not hold in more complex networks where a single shareholder may hold stock of several firms, as illustrated by the following example.

**Example 2.** The network contains 10 firms, as represented in Figure 2. Each of the firms 1, 2 and 3 holds 1/3 of the shares of firm 7. Each of the firms 4, 5 and 6 holds 1/3 of the shares of firm 8. Each of the firms 7, 8 and 9 holds 1/3 of the shares of firm $t = 10$. It is easy to see that, in each of the direct games $g_7$, $g_8$ and $g_{10}$ associated with firms 7, 8 and 10, all three shareholders have Banzhaf index equal to 0.5. In the compound game $v_{10}$, the sources have index $Z(i, 10) = 0.25$ for $i = 1, 2, 3, 4, 5, 6$, and $Z(9, 10) = 0.5$, in agreement with Eq. (6).

If we assume now that firm 3 merges with firm 4 (new name of the merged firm: 34), then the direct games $g_7$, $g_8$ and $g_{10}$ remain essentially unchanged, but the power indices of firms 34 and 9 in $v_{10}$ become: $Z(34, 10) = 0.5$ and $Z(9, 10) = 0.375$. Note, in particular, that the power index of firm 9 is affected in a rather unexpected way by the modification of apparently irrelevant shareholdings. □

![Figure 2: The network for Example 2](image)

Since our goal is to handle complex corporate networks involving hundreds or thousands of interlocking actors, we find ourselves – to some extent – in unchartered territory and we face a number
of difficult algorithmic issues. The purpose of this section is to explain the choices that we have made in order to resolve these issues. As a result of our choices, we end up computing a power index which, strictly speaking, is not the Banzhaf index but, as we shall argue, a very close substitute for it. In related studies, we have dubbed this number the Z-index or Zeno-index; see Crama, Kang, and Leruth (2004), Crama, Leruth, Renneboog, and Urbain (2003), Khatri, Leruth, and Piesse (2002).

3.1 Large networks

Let us assume that $G = (V, A; w)$ is an acyclic complete network, let $N$ be the set of sources of $G$, with $|N| = n$, and let $t \in V$ be an arbitrary target. In principle, the Banzhaf value $Z_G(j, t)$ can be computed (simultaneously) for all sources $j \in N$ by generating all possible $(2^n)$ vectors of votes $X$ for the source-players, by relying on Eq. (1) to propagate the votes down the network and to compute the outcome of the indirect game $v_t(X)$ for each vector $X$, and finally by using Eq. (5) to obtain $Z_G(j, t)$.

When the number of sources is very large, or when the computations must be carried out a large number of times, however, this approach quickly becomes inefficient (we shall see below that this is often the case for corporate networks). Therefore, we resort to Monte Carlo simulations in order to evaluate $Z_G(j, t)$. More precisely, we sample a set $S$ of vectors uniformly over $\{0, 1\}^n$ and we evaluate $v_t(X)$ for each $X \in S$. Note that the expression

$$\frac{1}{2^{n-1}} \sum_{X \in \{0, 1\}^n: x_j = 1} v_t(X)$$

appearing in Eq. (5) represents the average of $v_t$ over all vectors in $\{0, 1\}^n$ with $j$-th coordinate equal to 1. Similarly,

$$\frac{1}{2^{n-1}} \sum_{X \in \{0, 1\}^n: x_j = 1} v_t(X - e_j)$$

is the average of $v_t$ over all vectors with $j$-th coordinate equal to 0. Therefore, the quantity

$$\frac{1}{2^{|S|}} \cdot \left( \sum_{X \in S: x_j = 1} v_t(X) - \sum_{X \in S: x_j = 0} v_t(X) \right)$$

is an unbiased estimator of $Z_G(j, t)$.

3.2 Incomplete networks

When the network is not complete, we must specify a model for the float. Assume that the float holds a positive fraction $f l_j$ of the shares of $j$. Two limiting cases are described in Cubbin and Leech (1983) and Leech (1988, 2002a) (see also Nauenberg, Basu, and Chand (1997) for related
considerations in a different context). In the first case (concentrated float), each of the unidentified shareholders is assumed to own the same proportion of shares as the smallest identified shareholder: that is, each of them has weight equal to \( m_j = \min\{ w_{ij} : (i, j) \in A \} \) in the direct game \( g_j \) (with the possible exception of a smaller residual shareholder). In the second case (oceanic float), there is an infinite number of unknown shareholders and each of them holds a vanishingly small fraction of the shares (see Dubey and Shapley (1979), Milnor and Shapley (1978), Shapiro and Shapley (1978)). Dubey and Shapley (1979) have obtained limiting results for the Banzhaf index in oceanic weighted majority games, but their results do not seem to apply in our context.

In our computations, we simulate the float as a large number of small identical players, each with weight \( m_j \) (in practice, due to regulatory disclosure requirements, \( m_j \) usually takes a value around 0.03). Then, the total voting weight carried by the float is a random variable which can be approximated by a normal variable with mean \( fl_j / 2 \) and with variance \( fl_j \cdot m_j / 4 \).

### 3.3 Identification of controlling owners

In order to operationalize the concept of “ultimate owners” of a firm, La Porta, Lopez-De-Silanes, and Shleifer (1999) and other authors have proposed to consider owners of owners, and so on, while cutting off any ownership link below a certain arbitrary threshold (e.g., below 20%). Claessens, Djankov, Fan, and Lang (2002) also use a variety of shareholding cutoffs to identify the “ultimate owner” of a given firm. Climbing up the network in this fashion, however, may still produce a rather large number of small, unimportant players, as the chains of control are progressively diluted. We now explain how this analysis can be refined in the game-theoretic framework.

Consider an acyclic corporate network \( G \) and assume that we are interested in a specific target firm \( t \). As explained in Section 2.1, the game \( v_t \) associated with \( t \) is defined on the player-set \( N \), where \( N \) is the set of sources of \( G \). It is easy to see that the power \( Z_G(j, t) \) of a firm \( j \in N \) is always 0 when \( j \) is not an ancestor of \( t \). Therefore, in our computations, we can (and we do) restrict our attention to the subnetwork of \( G \) generated by the ancestors of \( t \).

Even so, some (and usually many) of the firms \( j \in N \) may have a very small power index \( Z_G(j, t) \), meaning that they have no significant influence on \( t \). When this is the case, we remove the irrelevant firms and we repeat the whole procedure on the resulting reduced network. Successive pruning of weak players eventually produces a residual subnetwork \( G_r^t \) with sink \( t \) and with source-set \( I(t) \), such that all firms \( j \in I(t) \) have relatively high power indices \( Z_{G_r^t}(j, t) \): we view \( I(t) \) as the set of ultimate, or controlling shareowners of \( t \).

Of course, it may happen that no shareholder of firm \( t \) has a high power index. When this happens, it is either because \( t \) has no identified shareholder at all (\( t \) may be an individual, or a company for which shareholding data is not available), or because the identified shareholders do not hold any substantial amount of control. In either case, we may consider that firm \( t \) is the head of its group and is controlled by its management. We return to this issue in Section 5.1 and in Section 5.3.
Figure 3: The residual subnetwork $G^r_5$

Figure 3 illustrates this discussion with a simplified example drawn from our analysis of a Korean database further described in Section 4. It provides a visual representation of the residual subnetwork $G^r_5$, consisting of firm $t = 5$ together with its controlling shareholder $I(5) = \{1\}$ and the set of firms lying on paths from firm 1 to firm 5 in $G$. Here, the threshold $q$ was set equal to 0.50.

### 3.4 Cycles

Real corporate networks are often cyclic, especially in Continental European and Asian markets featuring the insider system of governance; see e.g. Crama, Leruth, Renneboog, and Urbain (2003). Cycles raise a difficult issue in our approach, since they prevent the application of the recursive definition in Eq. (1). In other words, the game $v_t$ is ill-defined when the ancestors of vertex $t$ induce a cyclic network.

Gambarelli and Owen (1994) have made an attempt to provide a rigorous game-theoretic framework
for the structural analysis of corporate networks. Their approach relies on the concept of “consistent reduction” of the structure. Applying their results to the numerical computation of indices in real networks, however, is far from obvious, for several reasons: consistent reductions may not be uniquely defined, their computation - as presented by Gambarelli and Owen - requires an analytical multilinear expression of the game, and even so, the computation of a consistent reduction may be computationally expensive.

In order to account for these difficulties, we heuristically extend our approach as follows. We define a voting pattern to be an assignment of 0-1 values (such as arising in the course of a simulation run) to all vertices of the network. We say that a voting pattern $X \in \{0, 1\}^{|V|}$ is stable if, in the direct game $g_j$ associated with every individual firm $j$, the outcome of the vote is $x_j$ when every direct shareholder $i \in V_j$ votes $x_i$, i.e. if

$$x_j = g_j(X) \text{ for all } j \in V$$

(compare with Eq. (1); the notion of stability is closely related to the concept of “consistent reduction” in Gambarelli and Owen (1994)).

In the presence of cycles, we attempt to find a stable voting pattern, i.e. to solve the system of fixed-point equations (7) by successive iterations. In order to describe more precisely this procedure, let us assume that we are running a simulation as described in Section 3.1 and that the vote of the firms in $V^{fix} \subseteq V$ has already been fixed to the value $x_j^0$, for all $j \in V^{fix}$. Assume further that the procedure encounters a strong component $S \subseteq V$ whose votes are not entirely determined by the votes of its ancestors (see Example 3 below). We can choose $S$ in such a way that all its ancestors are in $V^{fix}$ (this is because the strong components of a graph can be labeled $S_1, S_2, \ldots, S_r$ so that there is no arc from $S_k$ to $S_\ell$ unless $k \leq \ell$; i.e., the strong components define a tree structure among themselves; see Bang-Jensen and Gutin (2000)).

**Example 3.** Consider the network in Figure 4. Assume that the vote of firms (or shareholders) 1 and 2 has been fixed to $x_1^0 = x_2^0 = 1$, while the vote of firm 3 has been fixed to $x_3^0 = 0$. Then, by propagation from vertices 1 and 2, the vote of firm 4 must be equal to 1, but all other votes remain undetermined. Thus, $V^{fix} = \{1, 2, 3, 4\}$ and we take $S$ to be the strong component $S = \{4, 5, 6, 7, 8\}$. \[\square\]
Let $\overline{S} = S \setminus V^{fix}$. Then, the following procedure $MIX$ is performed until a suitable termination criterion (to be described below) is fulfilled:

**Procedure MIX**

**(Initialization)** For each $j \in V^{fix} \cup \overline{S}$, the vote of $j$ is set to the value $x^0_j$, where:
- if $j \in V^{fix}$, then $x^0_j$ is the fixed vote of $j$;
- if $j \in \overline{S}$, then $x^0_j$ is the outcome prescribed by the direct game $g_j$ when every shareholder $i \in V_j \cap V^{fix}$ casts its fixed vote $x_i^0$ and the remaining shareholders in $V_j \setminus V^{fix}$ (as well as the float) vote randomly.

**(Repeat for $k = 1, 2, \ldots$)** For each $j \in V^{fix} \cup \overline{S}$, the vote of $j$ is set to the value $x^k_j$, where:
- if $j \in V^{fix}$, then $x^k_j = x^0_j$;
- if $j \in \overline{S}$, then $x^k_j$ is the outcome prescribed by the direct game $g_j$ when every shareholder $i \in V_j$ casts the vote $x_i^{k-1}$ and the float casts the vote $x_j^{k-1}$.

Let us formulate a few comments about this procedure (we shall come back to it in Section 6).

(a) Besides its interpretation as a solution method for Eq. (1), $MIX$ can also be viewed as reflecting the possible voting behavior of players in a cyclic network: at every moment in time, each firm observes the position adopted by its direct shareholders, then modifies its vote accordingly. This adjustment process repeats until a stable pattern of votes emerges.
(b) In the initialization step of MIX, each shareholder $i \in V_j \setminus V^{fix}$ votes randomly either 0 or 1. In our implementation, the probability of each vote depends on the amount of fixed votes already cast by the shareholders of $i$, i.e. it depends on the quantities
\[
\sum_{k \in V \cap V^{fix}} w_{ki} x_k^0 \quad \text{and} \quad \sum_{k \in V \cap V^{fix}} w_{ki} (1 - x_k^0).
\]

c) In all subsequent steps of MIX, the vote of the float of firm $j$ coincides with the tentative vote of $j$ recorded in the previous step. This is an attempt to stabilize the outcome of the procedure, by expressing the fact that the vote of firm $j$ in the current iteration should depend not only on the vote of the other shareholders, but also on its own vote in the previous iteration. One possible “behavioral” interpretation is that the float is identified with, or is influenced by, the management of the firm, which may have its own independent position on the issues at stake.

d) Procedure MIX is not guaranteed to converge to a stable pattern. However, since its iterations are deterministic (except the first one) and since the number of distinct voting patterns is finite, the sequence $(X^k; k \geq 0)$ generated by MIX is necessarily periodic (this simple observation would not apply if the float were to vote randomly at every iteration). But the period of this sequence may be quite long, as illustrated by Example 4 hereunder. Therefore, in our implementation, we stop the procedure when $X^k$ is “sufficiently close” to $X^{k-1}$, as measured by the number of firms which do not change their vote from step $k - 1$ to step $k$, or when $k$ grows too large.

Example 4. Consider a directed cycle of $n$ firms where each firm is fully owned by a unique shareholder, and suppose that $V^{fix} = \emptyset$. If all firms initially vote 1, except firm $n$ which votes 0, then MIX goes through a sequence of $n$ successive iterations before it reproduces the initial pattern.

3.5 Preprocessing

The algorithmic ideas described in previous sections have been implemented and extensively tested on numerous real-world databases, with sizes ranging from a few dozen to more than 2000 arcs. In order to be able to handle large databases, special attention has been devoted to the computational efficiency of the implementations. In particular, we found it useful to preprocess the databases in the following way. Suppose that a given firm, say firm $i$, surely controls another firm $j$ in the sense that $i$ owns more than 50% of the shares of $j$ (the terminology “sure control” is borrowed from Salvemini, Simeone, and Succi (1995)). Then, we can delete firm $j$ from $G$ while simultaneously discarding all the arcs of the form $(\ell, j)$ and replacing each arc of the form $(j, k)$ by an arc $(i, k)$ with value $w_{jk}$. Indeed, since the vote of firm $i$ is always transferred integrally to firm $j$, this transformation has no effect on any voting pattern, nor, a fortiori, on the power index of any firm.
The concept of “sure control” can be recursively extended to more complex pyramidal situations as follows: if \( i \) surely controls a subset \( S \) of direct shareholders of \( j \) and if \( \sum_{k \in S} w_{kj} > 0.5 \), then we also say that \( i \) surely controls \( j \) (indeed, \( i \) owns more than 50% of \( j \) in the reduced graph obtained after preprocessing all the arcs \((i, k), k \in S\)).

Preprocessing was found to reduce the size of most databases in a very significant way and to accelerate accordingly the computation of power indices.

4 Numerical validation

We have performed several numerical tests in order to demonstrate the stability of our procedure. We report here on two such tests.

First, we consider the weighted majority game corresponding to the votes of all 51 member-states in the U.S. Presidential Electoral College (we rely on the figures of the 1990 Census data as found in Leech (2003)). We have computed the power index of each state 10 times, by performing 10 independent runs of our algorithm. Table 1 shows the results obtained for 15 representative states of various sizes. Column 1 identifies each state, column 2 contains the true value \( Z \) of its power index, as given by Leech (2003), column 3 displays the average value \( \bar{Z} \) of its power index for 10 runs of our algorithm, column 4 displays the standard deviation \( \sigma \) of these 10 runs, and column 5 displays\(^4\) the normalized difference \( |(\bar{Z} - Z)/(\sigma/\sqrt{10})| \).

<table>
<thead>
<tr>
<th>State</th>
<th>True index</th>
<th>Average value</th>
<th>Standard deviation</th>
<th>Normalized difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>0.0229</td>
<td>0.0233</td>
<td>0.005892</td>
<td>0.261</td>
</tr>
<tr>
<td>AL</td>
<td>0.0687</td>
<td>0.0743</td>
<td>0.012734</td>
<td>1.381</td>
</tr>
<tr>
<td>AR</td>
<td>0.0457</td>
<td>0.0442</td>
<td>0.005237</td>
<td>0.931</td>
</tr>
<tr>
<td>AZ</td>
<td>0.0610</td>
<td>0.0648</td>
<td>0.010803</td>
<td>1.096</td>
</tr>
<tr>
<td>CA</td>
<td>0.4665</td>
<td>0.4639</td>
<td>0.008398</td>
<td>0.977</td>
</tr>
<tr>
<td>CO</td>
<td>0.0610</td>
<td>0.0592</td>
<td>0.009990</td>
<td>0.569</td>
</tr>
<tr>
<td>CT</td>
<td>0.0610</td>
<td>0.0607</td>
<td>0.009750</td>
<td>0.096</td>
</tr>
<tr>
<td>DC</td>
<td>0.0229</td>
<td>0.0205</td>
<td>0.007102</td>
<td>1.030</td>
</tr>
<tr>
<td>DE</td>
<td>0.0229</td>
<td>0.0207</td>
<td>0.008734</td>
<td>0.780</td>
</tr>
<tr>
<td>FL</td>
<td>0.1939</td>
<td>0.1951</td>
<td>0.007473</td>
<td>0.514</td>
</tr>
<tr>
<td>MO</td>
<td>0.0841</td>
<td>0.0864</td>
<td>0.009395</td>
<td>0.781</td>
</tr>
<tr>
<td>NC</td>
<td>0.1072</td>
<td>0.1104</td>
<td>0.007210</td>
<td>1.420</td>
</tr>
<tr>
<td>NY</td>
<td>0.2596</td>
<td>0.2573</td>
<td>0.014694</td>
<td>0.488</td>
</tr>
<tr>
<td>OH</td>
<td>0.1620</td>
<td>0.1655</td>
<td>0.010087</td>
<td>1.117</td>
</tr>
<tr>
<td>TX</td>
<td>0.2512</td>
<td>0.2496</td>
<td>0.010483</td>
<td>0.493</td>
</tr>
</tbody>
</table>

Table 1: Power indices for the U.S. Electoral College

The difference between the true value of the power index and the average value over 10 runs is

\(^4\)The values in column 5 cannot be retrieved exactly from the previous columns, due to round-off errors.
never significant, for any of the states\(^5\). As a matter of fact, the values \(Z\) and \(\overline{Z}\) are always within 0.006 of each other: we believe that this level of accuracy is largely sufficient for most practical applications of power indices.

The second test relies on a large database of the shareholding structures around firms listed on the Seoul Stock Exchange (the same database has been used by Crama, Kang, and Leruth (2004)). The database contains the names of about 2080 firms; 667 of these firms have at least one identified shareholder, for a total of 2240 shareholding links which create a complex network of cyclic relationships.

For each firm \(t = 1, \ldots, 667\) with explicitly identified shareholders, we have estimated the power \(Z^1(j, t)\) of each firm \(j\) in the indirect games \(v_t\). Then, for each \(t\), we have recorded the name of the firm with the largest value of the power index, i.e. the firm \(j_t\) such that \(Z^1(j_t, t) = \max_i Z^1(i, t)\). We have repeated the same procedure five times, thus getting five independent estimates of \(Z(j_t, t)\), say \((Z^k(j_t, t) : k = 1, \ldots, 5)\) for each firm \(t\).

Since we do not know the “real” value of \(Z(j_t, t)\), our main goal in this experiment was to assess the robustness of the results over independent runs of our algorithms. Indeed, the existence of large pyramidal structures and of cycles might be expected to increase the variance of the estimates. In particular, it is not even obvious that the “most influential” firm \(j_t\) will remain unchanged in successive runs \(k = 1, \ldots, 5\). In our experiments, however, \(j_t\) remains constant from run to run for 627 out of the 667 firms. For the remaining 40 targets \(t\), \(Z^k(j_t, t)\) was not always the largest value of the power index, but it was always extremely close to it. In other words, \(j_t\) was not systematically recognized as the most influential firm for all \(t\), but confusion arose only in those cases where there was a tie or a very close competition for the award.

For all firms \(t\), we have computed the average \(M(t)\), the range \(R(t)\), and the relative range \(RR(t) = R(t)/M(t)\) of \(Z^k(j_t, t)\) over the five runs, where \(R(t) = \max_k Z^k(j_t, t) - \min_k Z^k(j_t, t)\). The range \(R(t)\) is always quite small: indeed, its average value (over \(t = 1, \ldots, 667\)) is 0.02, and its maximum value is 0.08. The average value of the relative range \(RR(t)\) is also small, at 0.05. However, \(RR(t)\) may occasionally assume rather large values, up to 0.60. This happens very rarely and only when the largest power index is very small: indeed, when restricted to those (410) firms with \(Z^1(j_t, t) \geq 0.5\), \(RR(t)\) averages 0.025 and never exceeds 0.11.

From these tests, and from other experiments not reported here, we conclude that the algorithms described in previous sections are reasonably robust and can be meaningfully used to analyze large and complex shareholding databases.

\(^5\) The critical value of the Student \(t\) statistic, for \(\alpha = 0.10\) and with 9 degrees of freedom, is 1.83.
5 Applications

Several applications of Banzhaf indices have already been mentioned in previous sections. We are now going to sketch a number of additional applications that have been carried out using the concepts and algorithms presented above.

5.1 Identification of corporate groups

In Section 3.3, we have explained how the controlling shareholders of a given firm $t$ can be identified by repeated computations of the power indices in a sequence of reduced subnetworks. The procedure eventually returns a set $I(t)$ of controlling shareholders whose power indices $Z(j, t)$ ($j \in I(t)$) lie above a predetermined threshold $q$. If we perform the same procedure for all possible targets $t$, we eventually identify a collection $I = \bigcup_t I(t)$ of firms that can be regarded as highly influential in the network $G$ (remember that the qualifier “influential” refers here to the a priori ability of the firm to exert control; whether or not the firm decides to actually exert this control may be a matter of strategy or opportunity).

For each $j \in I$, we can now define the set of firms $\{ t : Z(j, t) \geq q \}$ to be the group controlled by firm $j$. Thus, this approach allows us to automatically identify the various subgroups composing a large corporate network. Note that there may exist shareholding relationships between different subgroups. But if such relationships exist, they are not associated with any significant level of control.

As an example, Figure 5 maps the group controlled by firm $j = 1$ in the Korean database (compare with Figure 3).

5.2 Control and financial performance

The relation between ownership structure and financial performance has been investigated by numerous researchers; see e.g. Jensen and Meckling (1976), Shleifer and Vishny (1997), Claessens et al. (2000, 2002), Khatri, Leruth, and Piesse (2002), etc.

We reproduce here some of the results of Crama et al. (2003), who investigated the possible link between the performance of a stock and the ownership structure of the firm at the previous period, based on 1988-1993 data collected for a random sample of 250 companies quoted on the London Stock Exchange.

Several regression models have been tested by Crama et al. (2003). The dependent variable is the annual return on shares, which is selected as a proxy for financial performance. The independent variables are the level of employment, market capitalization, price-earning ratio, systematic risk indicator (Beta), and sector of activity of each firm, as well as different combinations of four additional variables meant to capture the ownership structure of each firm, namely:
Figure 5: The group controlled by firm 1

(a) $Z_1$, the power index of the largest shareholder;
(b) $Z_2$, the power index of the second largest shareholder;
(c) a Herfindahl index of concentration $H_1$, defined as the square of the percentage of shares held by the largest shareholder;
(d) another Herfindahl index $H_5$, defined as the sum of squared percentages of shares owned by the five largest shareholders.

The variables $H_1$ and $H_5$ have been included because they capture, in a simple way, the link between larger shareholdings and higher levels of control. By contrast, the inclusion of $Z_1$ and $Z_2$ is fully consistent with the approach taken in this paper. Three different regression models therefore include either $H_1$ and $H_2$, or $Z_1$ and $Z_2$, or $Z_1$ only, respectively.
Interestingly, the coefficients of $Z_1$ and $Z_2$ are highly significant in both models where they appear. These conclusions support the claim that Banzhaf indices adequately capture the relationship between shareholdings and control, and thus provide a useful instrument for modeling ownership structure. By contrast, the coefficients of $H_1$ and $H_5$ are only significant at a very low level in all models. Thus, the results show that, even in the simple case of an outsider system (of which the British equity market is an archetype), the $Z$-index performs better than the Herfindahl index.

In another paper, Khatri, Leruth and Piesse (2002) have checked that similar conclusions apply to more intricate structures. They analyze the performance of 40 firms listed on the Kuala Lumpur Stock Exchange, which provides a typical example of an insider system. The corresponding shareholder database contains more than 1200 firms and 1100 links, and includes a few cyclic substructures. Here again, an econometric analysis based on stochastic production frontiers underlines the usefulness of the $Z$-index as a proxy of ownership concentration.

5.3 Market liquidity

In Crama, Kang, and Leruth (2004), the basic premise is that shares associated with significant levels of control are less liquid than the remaining ones. This premise is used to derive a measure of liquidity for a stock and, by extension, for a financial market: essentially, the *liquidity* of a stock is defined as the fraction of the market value of its shares which is not significantly controlled. We refer to Crama, Kang, and Leruth (2004) for a more precise definition as well as for an application to the Korean stock exchange, and we simply provide here an informal discussion of the concept.

To take an extreme example, consider a target firm $t$ and assume that one of its controlling shareholders $j$ has power index equal to 1 (say, $j$ owns 60% of the shares of $t$). Then, the target is fully controlled by shareholder $j$. One can thus consider that while the shares controlled by this main shareholder are not liquid (in the sense that their value is higher than the market price of the float, because it reflects the control premium), the remaining shares could be considered as liquid at the prevailing price.

At the other extreme, consider a situation where the target has no controlling shareholder. In that case, we have indicated that the target is assimilated to the head of a group and can be viewed as controlled by its management (if the target is a firm) or as an individual shareholder. If it is a firm, we can consider that all its shares are liquid.

Finally, in an intermediate situation, there may be one important controlling shareholder for $t$ with power index close to 1, but strictly smaller than 1. Note that this main shareholder may own substantially less than 50% of the shares, for example if the float is large or if there is a large number of dispersed shareholders. The target will then usually have one or two additional main shareholders with low power indices. The question is now to select a suitable threshold for the $Z$-index, below which the level of control is too low to really matter. Those shares associated with levels of control below the threshold are considered liquid. One important aspect of the study by
Crama, Kang, and Leruth (2004) is to analyze the change in liquidity ratios as the threshold varies. To illustrate how it is done, consider the following example.

**Example 5.** There are two targets $A$ and $B$. For each target, 100 shares are distributed among some shareholders or belong to the float. The share price is normalized to 1 without loss of generality. Shareholder 1 owns 56 shares of $A$ and Shareholder 2 owns 7 shares of $A$. The float for $A$ consists of 37 shares. Shareholder 3 owns 13 shares of $B$, Shareholder 4 owns 7 shares of $B$, and Shareholder 5 owns 4 shares of $B$, while 72 shares constitute the float for $B$.

The value of the $Z$-indices and the liquidity computations are summarized in Table 2. We see that, at the threshold 0.50, 44% of the shares of $A$ and 87% of the shares of $B$ are liquid, for a total market liquidity of 65.5%. At the threshold 0.25, 44% of the shares of $A$ are still liquid, but only 80% of the shares of $B$ can be regarded as liquid, resulting in a total market liquidity of 61%. In either case, the liquidity estimates are higher than those which would be obtained by focusing solely on the value of the float.

<table>
<thead>
<tr>
<th>Shareholder</th>
<th>Shares</th>
<th>$Z$-index</th>
<th>Liquid at threshold 0.50</th>
<th>Liquid at threshold 0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sh1</td>
<td>$w_{1,A} = 0.56$</td>
<td>$Z(1, A) = 1$</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sh2</td>
<td>$w_{2,A} = 0.07$</td>
<td>$Z(2, A) = 0$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Float</td>
<td>$fl_A = 0.37$</td>
<td>$Z(fl, A) = 0$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sh3</td>
<td>$w_{3,B} = 0.13$</td>
<td>$Z(3, B) = 0.55$</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sh4</td>
<td>$w_{4,B} = 0.07$</td>
<td>$Z(4, B) = 0.25$</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Sh5</td>
<td>$w_{5,B} = 0.04$</td>
<td>$Z(5, B) = 0.14$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Float</td>
<td>$fl_B = 0.76$</td>
<td>$Z(fl, B) = 0$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2: Power indices and liquidity at various thresholds

In Crama, Kang, and Leruth (2004), significant differences are similarly found between the total value of floating shares and the value of liquid (up-for-grab) shares on the Koran stock exchange.

While we are not aware of any other methodology to systematically measure liquidity (other than that used by large financial data providers who rely on ad hoc classifications to determine whether shares are liquid or not), we should nevertheless mention some weaknesses of our approach.

First, some shareholders may not be willing to sell their shares, even if they hold no control, unless the price offered substantially exceeds the market price (say, a speculator who expects the price to climb even higher, or a free-rider as in Grossman and Hart (1980)): liquidity depends on how large a margin a potential buyer is willing to offer over the prevailing market price.

Second, the major shareholders (with power indices above a given threshold) usually own shares in excess of the amount required to reach the control threshold (for instance, in Example 5, Shareholder 1 only needs 50.1% of the shares in order to hold full control). In such situations, some of the shares of the major shareholder could be sold without giving up control, and hence could possibly be viewed as “liquid” as well. Such considerations have not been incorporated in our model: they implicitly
refer to the distinction between “control” and “ownership”, as well as to the principal-agent issues which have been briefly mentioned in the introduction of this paper and, as such, go well beyond the scope of the present study.

Finally, we should point that liquidity is also influenced by many other factors, such as trading restrictions or tax regulations. Our model does not capture these aspects of liquidity.

6 Conclusions and directions for further research.

In this paper, we have proposed an algorithmic approach for the computation of power indices associated with corporate networks. Our approach to the measure of control is based on a formal game-theoretic framework. We have focused our analysis on the computation of Banzhaf indices. Since computing the exact value of the indices is a hard algorithmic problem, we have proposed Monte Carlo simulation techniques to obtain fast approximations. The difficulties associated with these techniques in very large networks have been identified. We have demonstrated their practical feasibility on several large-scale databases, and we have described several economic studies relying on our methodology.

A major question for future research would be to develop a rigorous mathematical model which would properly account for the occurrence of cycles in real-world corporate networks (our current procedure for dealing with cycles is mostly heuristic). Elements of solutions to this question can possibly be found in the work of Gambarelli and Owen (1994) (see Section 3.4), in the literature on stochastic games or on dynamical systems. The neural network literature, for instance, contains results about the periodic behavior of undirected networks of majority gates, but these results do not extend immediately to directed networks; see e.g. Goles and Tchuente (1983), Poljak and Sura (1983). It may also prove interesting to investigate the applicability, in our context, of a model recently proposed by Hu and Shapley (2003) in order to derive an equilibrium authority distribution for organizations.

Even tough cycles may eventually disappear from shareholding structures, they also exist in other types of financial networks. For instance, we believe that our work could open the way for further applications of power indices to the analysis of Real Time Gross Settlement agencies, where members take an insurance against the possibility of default by other members. These insurances tend to be very costly and the level of protection is high (essentially, the RTGS will pay on behalf of the defaulter), as is the level of punishment (members are not required to pay their dues to the member who defaulted until it clears the debt). Measuring the actual risk (and hence computing the level of the premium) would possibly be facilitated by using an approach similar to that developed in this paper.
References


