

Global ownership and corporate control networks

Armando Rungi[§]
(IMT Lucca)

Gregory Morrison[†]
(University of Houston)

Fabio Pammolli*
(Politecnico di Milano)

July 2017

Abstract

In this contribution, at first, we introduce a basic network framework to study pyramidal structures and wedges between ownership and control of companies. Then, we apply it to a dataset of 53.5 million of companies operating in 208 countries. Among others, we detect a strong concentration of corporate power, as less than 1% of parent companies collect more than 100 subsidiaries, but they are responsible for more than 50% of global sales. Therefore, we show that the role of indirect control, i.e., through *middlemen subsidiaries*, is relevant in 15% of domestic and 54% of foreign subsidiaries. Among foreign companies, cases emerge of blurring nationality, when control paths cross more than one national border, in the presence of *multiple passports* (19.1%), *indirectly foreign* (24.5%), and *round-tripping* subsidiaries (1.33%). Finally, we relate indirect control strategies to country indicators of the institutional environment. We find that pyramidal structures arise less likely in the presence of good financial and contractual institutions in the parent's country, as these foster more transparent forms of corporate governance. Instead, parent companies choose indirect control through countries of subsidiaries that have better financial institutions, possibly because it is easier to coordinate decisions from remote. Finally, we find that offshore financial centers are preferred jurisdictions for *middlemen* subsidiaries, probably due to a lower taxation and a lack of financial disclosure.

Keywords: ownership, corporate control, corporate governance, financial networks, financial institutions, offshore, economic entrenchment

JEL codes: G32; G34; F36; F23; C63; C71; L14

A previous version of this paper circulated with the title 'Corporate boundaries of Multinational Enterprises: a network approach'. We would like to thank the staff of the Division on Investment and Enterprise of the United Nations Conference on Trade and Development (UNCTAD) for fruitful comments during Expert Group Meetings organized while drafting the World Investment Report 2016. We also thank participants to the first workshop on 'Networks, Complexity, and Economic Development', organized by the Hungarian Academy of Sciences, and participants to the European Trade Study Group (ETSG 2015) at Université Paris 1 - Panthéon-Sorbonne. The authors claim financial support by CrisisLab PNR Project at IMT School for Advanced Studies, Lucca.

[§]Corresponding author: mail to armando.rungi@imtlucca.it. Laboratory for the Analysis of Complex Economic Systems, IMT School for Advanced Studies, piazza San Francesco 19 - 55100 Lucca, Italy.

[†]mail to: gcmorris@central.uh.edu. Department of Physics. University of Houston, TX, USA

*mail to: fabio.pammolli@polimi.it. Dipartimento Ingegneria Gestionale. Politecnico di Milano, Milano, Italy

1 Introduction

In 2016, according to the Fortune magazine’s ranking, top 500 global companies employed about 67 million people worldwide, generating 27.7 trillion US dollars of consolidated revenues. In the same year, cross-border mergers and takeovers amounted to 869 trillion US dollars, showing a rising trend over the latest five years (UNCTAD, 2017). Despite different national frameworks for corporate governance, the ownership of firms and the market for corporate control are by now global activities, with a propensity to shift from models of dispersed ownership to models of concentrated ownership (OECD, 2017).

In fact, a governance based on concentrated ownership contradicts the epitome of a Modern Corporation, once sketched in the seminal work by Berle and Means (1932; 1967), with widely held capital shares and small shareholders relying on accountable professional managers. Nothing new under the sun, since La Porta et al. (1999) have shown how one or more shareholders can have voting rights significantly in excess of cash flow rights, mainly through indirect control over pyramidal ownership structures.

More recently, Levy and Szafarz (2017) studied how cross-ownership linkages may also allow managers (not shareholders) to seize indirect voting rights, hence insulating their firms from outside shareholders’ control. They discuss how cross-ownership is nowadays a phenomenon that can be found both in civil law countries and, for example, in the US mutual fund industry.

Whether indirect voting rights are seized by ultimate shareholders or by managers, pyramidal structures and cross-ownership pave the way for a concentration of economic power by few, who do not make a commensurate capital investment, hence endangering innovation rates, allocation of financial resources, and economic growth (Morck et al., 2005).

Briefly, ownership and corporate control may be separated. The phenomenon has been studied both theoretically and empirically (Mathews, 2007; Almeida and Wolfenzon, 2006; Faccio and Lang, 2002; Claessens et al., 2000; La Porta et al., 1999; Fama and Jensen, 1983), often concerning the institutional environments of the origin countries. However, previous works often relied on few albeit representative case studies, as in La Porta et al. (1999), or on a geographic subset of quoted firms, as in Faccio and Lang (2002). Thankfully, an increasing availability of data on shareholding activity allows for a more systematic study of corporate control structures on a global scale.

In this contribution, we first propose a simple network perspective to tackle with the separation of ownership and control of companies, which is flexible enough to describe the design of different corporate governance structures around the world. Therefore, we apply our network framework to a unique dataset of about 53.5 million of companies active in 208 countries in 2015. In this way, we can provide for the first time an overall assessment of indirect control through fragmentation/concentration of voting rights, within and across national borders. Finally, we empirically relate the emergence of indirect control to the quality of country level institutions, controlling for political and physical geographic characteristics of origin and destination countries of investment operations.

We build on previous literature when we introduce a theoretical framework to separate corporate control from fragmented ownership (Levy and Szafarz, 2017; Levy, 2011; Chapelle and Szafarz, 2007; Crama and Leruth, 2007), which considers the concentration, the transitivity and the consolidation of voting rights in webs of interlocking assemblies of shareholders. However, our purpose is to extract corporate control structures from raw shareholding data, whereas previous works intended to study the corporate power of agents within already assigned control structures.

As far as we know, ours is the first attempt to apply a big data algorithm to global shareholding activities. Besides computational obstacles, our effort required the introduction of a general ownership space, which we imagine as a common playfield for every possible investor in the share capital of firms, whose original goal is either to participate in the distribution of dividends and to exert influence on companies’ management decisions. In this context, we assume that any corporate structure embeds in the ownership space, as it can be depicted in the form of a peculiar graph, i.e., a hierarchy of firms, made of a parent and its subsidiaries. In these hierarchical graphs, a parent can coordinate the economic activity of its subsidiaries, possibly located on different layers of control distance from the parent, through direct or indirect equity linkages. It is in this context that we introduce a notion of

corporate control distance as the number of intermediate *middlemen* subsidiaries that are needed to enforce decisions from headquarters.

In a second part, when we extract corporate control from global ownership, we unveil hidden heterogeneity in the design of corporate structures. First, we detect a strong concentration of economic power by few parent companies, because only less than 1% of the total collect more than 100 subsidiaries, but these are responsible for more than 50% of global sales in our data. Then, we find that indirect corporate control is more common than we expected, because 'subsidiaries of subsidiaries' represent about 15% of domestic and 54% of foreign companies.

More specifically, peculiar cases emerge among foreign affiliates, which we document for their policy relevance in the regulation of competition, international investment, and taxation. We find that the classification of companies by country of origin¹ can be ambiguous:

i) because a company can have *multiple passports*, when indirect control paths cross multiple countries, in 19.1% of our data on foreign affiliates;

ii) because a company can be only apparently domestic, after we look at the immediate shareholders, but it proves to be *indirect foreign*, once we check for upper layers in indirect control paths, in 24.5% of our data on foreign affiliates;

iii) because a domestic investor can use a foreign country to reinvest back in her country of origin, hence starting a so-called *round-tripping* investment, possibly to exploit the advantages of a foreign jurisdiction, in a non-negligible 1.33% of our data on foreign affiliates.

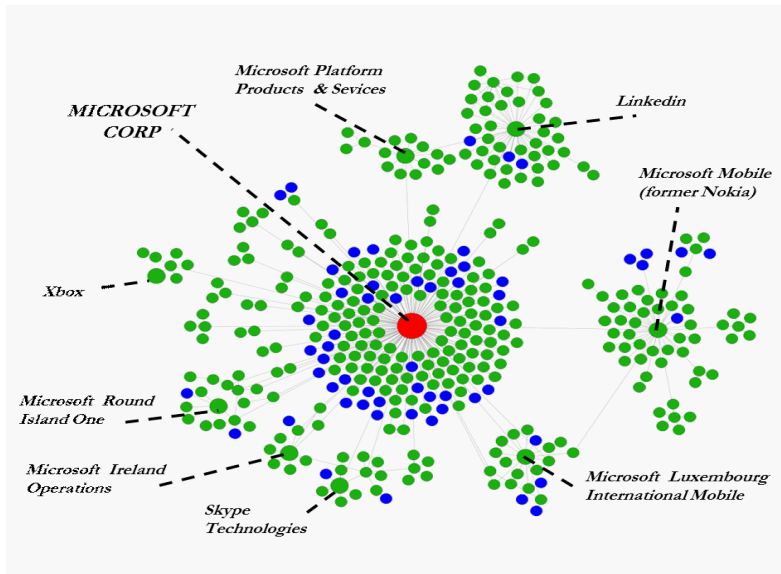
In a third part, we relate indirect control to indicators of geography, contractual and financial frictions at the country level. Indirect control is the way pyramids and vertical structures arise. We find that financial institutions in the countries of origin and destination of investment are indeed drivers of indirect control, but with an opposite sign. In particular, we find that parent companies have fewer pyramidal structures when their countries have better financial institutions, since the latter foster more transparent forms of corporate governance. However, *ceteris paribus*, better financial institutions in both the countries of *middlemen* and final subsidiaries increase the probability that a parent company establishes indirect control. The latter finding is coherent with the idea that, once vertical structures are established, their operability requires lower frictions along control paths so that management decisions can be enforced from the top of headquarters.

Finally, we detect a role for offshore financial centers, as preferred jurisdictions for *middlemen* subsidiaries, possibly due to their lower levels of taxation and a lack of transparency in financial disclosure.

To have an idea of the complexity of corporate networks we may encounter, we show two examples extracted from our elaborations in Figure 1 and Figure 2. Microsoft Corp and the Toyota Group are among the top 0.1% corporate networks that we detect, regarding both numbers of subsidiaries and consolidated sales. They are also included in the ranking by Fortune Global 500. They show different corporate structures, and they were born at different times. The first is a leading IT company in the provision of computer software and consumer electronics founded in 1975. According to our elaborations, it collects 389 subsidiaries operating in 79 countries. The second is one of the biggest manufacturers of motor vehicles historically organized as a Japanese keiretsu, i.e., as a conglomerate structure with cross-shareholding exchanges among constituent firms. It was born in the 1930s, and according to our elaborations, it collects 2,239 subsidiaries organized on top by eight different parent companies with subsidiaries active in 125 countries. Each node in the graphs of Figure 1 and Figure 2 represents a legally autonomous firm that participates or is participated by other companies. Individual shareholders are excluded from visualization, although they are considered when drawing the boundaries of corporate control. The participations to share capital are depicted as edges between firms. In the case of Microsoft, most of the ownership linkages run from the parent company to subsidiaries, i.e., a majority of them is directly controlled, although some organize in clusters

¹See also UNCTAD (2016) for a discussion on the relevance of these cases, while adopting our procedure with similar data. More in general, UNCTAD (2016) usefully discusses the policy challenges set by complex ownership and control structures, when international investment agreements try to discriminate based on the nationality of investors.

Figure 1: The corporate control network of Microsoft



Corporate network of Microsoft Corp visualized with a ForceAtlas2 layout (Jacomi et al., 2014) using Gephi software. A red node indicates a parent company, while green nodes represent majority-owned subsidiaries and blue nodes are affiliates with a dominant stake.

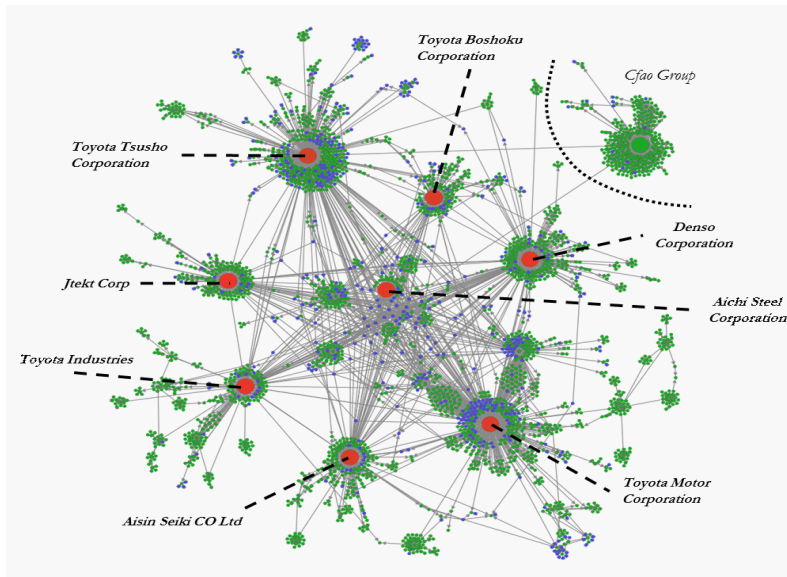
around LinkedIn, Microsoft Mobile (former Nokia) and Skype Technologies. Indeed, these are all former parent companies acquired over time by Microsoft together with their subsidiaries, therefore embedded as such within the corporate hierarchy. Other clusters can be detected around Microsoft Ireland Operations, Microsoft Round Island One, Microsoft Luxembourg International Mobile, Xbox and Microsoft Platform Products and Services, which instead are all born inside Microsoft as financial or technological holdings.

In 2, the Toyota Group has a more sophisticated corporate structure with patterns of ownership often crossing each other, as we can expect in a Japanese *keiretsu*, where cross-holdings are a signal of long-term reciprocal commitment among constituent firms because managers can sit at interlocking assemblies. Exceptionally, the Toyota Group has eight different parent companies on top of its corporate structure, each of them holds a minority albeit dominant stake in the other seven. The latter seems not to be a standard solution in our elaborations since the greatest majority of companies prefer to organize with just one parent on top. A separate cluster of firms is detected around CFAO, which indeed was an autonomous group acquired by Toyota only in 2012, specialized in the distribution of cars, consumer goods, and pharmaceutical products in African countries.

What both these corporate structures have in common is the exercise of indirect control on some subsidiaries, which is pivotal in the design of pyramids and vertical structures. That is, some subsidiaries are at a proprietary variable distance from headquarters. This is the reason why we focus on the emergence of indirect control in our empirical investigations, as a starting point to approach the complexity of these objects.

The rest of the paper is organized as follows. Section 2 introduces our network framework for ownership of companies. In Section 3 we describe the identification of corporate control networks as a result of the concentration of voting rights on ownership networks. Section 4 describes data and introduces first empirical evidence. In Section 5 we discuss the empirical strategy and related results.

Figure 2: The corporate control network of the Toyota Group



Corporate network of the Toyota Group visualized with a ForceAtlas2 (Jacomi et al., 2014) layout using Gephi software. Red nodes indicate (interlocked) parent companies, while green nodes represent majority-owned subsidiaries and blue nodes are affiliates with a dominant stake in equity.

2 An ownership space

2.1 Ownership networks

In this section we introduce the notion of ownership space as developed by firms and their shareholders. Ownership networks emerge because: i) companies can hold stakes in other companies and become corporate shareholders; ii) stakes of companies can be fragmented among different (individual or corporate) shareholders; iii) a (corporate or individual) shareholder can have a stake in more than one company.

We can think of the ownership space as the common playfield of all investors with a stake in the share capital of firms, whose original goal is either to participate to the distribution of dividends and/or to exert influence on companies' management decisions.

Let define an ownership space, Ω , as a collection of ownership networks \mathcal{G} 's as follows:

$$\mathcal{G}(N_{\mathcal{G}}, V_{\mathcal{G}}, K_{\mathcal{G}}, W_{\mathcal{G}}) \in \Omega(N, V, K, W) \quad (1)$$

Each ownership network is made of: i) a set of nodes, $N_{\mathcal{G}}$, which are either companies or their shareholders; ii) a set of equity values for nodes, $V_{\mathcal{G}}$; iii) a set of ownership linkages running from shareholders to companies, $K_{\mathcal{G}}$; iv) a set of ownership weights $W_{\mathcal{G}}$ (as percentages) for each ownership linkage. The ownership space can be made of more than one ownership network, such that $\bigcup_g \mathcal{G} = \Omega$, where g is the number of separate networks. Ownership networks may differ along any of the dimensions in eq. 1.

Among nodes, N , we can find both companies ($I \subset N$) and shareholders ($S \subseteq N$). A company can also become a (corporate) shareholder², such that $I \cap S \neq \emptyset$.

We identify a subset of shareholders as *ultimate owners* ($U \subset S$). These are the ultimate beneficiaries of ownership, which directly or indirectly benefit from the productive activities of companies after the distribution of profits (or losses). Among them, we can find households or public authorities (governmental agencies, regional agencies, states). By nature and human institutions, the *ultimate owners* cannot be owned by any other entity, i.e. they are the original *sources* of any ownership network. That is, any set $U_{\mathcal{G}} \subset N_{\mathcal{G}}$ and $U_{\mathcal{G}} \neq \emptyset$.

Each company $i \in I$ has at least one *direct shareholder* s_i sitting in the corporate assembly S_i . If the ownership of a company is fragmented, we have more than one *direct shareholder*, $s_i \in S_i$, such that the corporate assembly is a subset of the total number of nodes, $S_i \subseteq S \subseteq N$, with cardinality $|S_i| \in [1, N]$.

The number of shareholders $|S_i|$ of the company i represents also its *ownership in-degree*, i.e. it counts how many *ownership linkages* end in the i th company.

The number of stakes held by each shareholder represents its *ownership out-degree*, i.e. it counts how many *ownership linkages* start from s_i and end in any company, including company i .

We can now define an *ownership relationship*, $\mathcal{S} : S \rightarrow I$, such that an *ownership linkage* ($k \in K$) runs from any owner/shareholder s_i to a company i . The ownership relationship is always *directed*, from the shareholder to the company, since the first holds voting rights and participates to dividends of the latter. The amount of stake a shareholder can hold in a firm falls in a percent range $w_{s_i i} \in W = (0, 100]$. In fact, any $w_{s_i i}$ represents the *ownership weight* attributed to any *ownership linkage*.

In turn, ownership rights are based on a distribution of equity³ at the company-level ($v_i \in V_i$), such that $\sum_i v_i = E$ is the total amount of equity scattered throughout the global economy at a given moment. We can further assume that *ultimate owners* have an equity value always equal to zero, since they cannot be owned.

Eventually, the ownership space Ω is a perfect partition of all ownership networks \mathcal{G} 's, such that $\bigcup_{\mathcal{G}} N_{\mathcal{G}} = N$, and $\bigcup_{\mathcal{G}} K_{\mathcal{G}} = K$, but also $\bigcap_{\mathcal{G}} N_{\mathcal{G}} = \emptyset$ and $\bigcap_{\mathcal{G}} K_{\mathcal{G}} = \emptyset$.

It means that in a world where ownership can be fragmented, a company does not participate to more than one ownership network and a shareholder does not own stakes outside of its network.

Nonetheless, among the possibilities that can be realized in an ownership space, we do not exclude the case of a unique ownership network, when all its components are possibly connected in a unique network.

On the other hand, a single ownership network \mathcal{G} may not be a partition on single corporate assemblies. Indeed, $\bigcup_i S_i$ has to sum up to the total number of shareholders in the ownership network, but we may have $\bigcap_i S_i \neq \emptyset$. That is, a single shareholder (individual, state or company) can participate to the equity of more than one firm.⁴

2.2 Ownership paths and ownership distance

We can now define a notion of path and a notion of distance in the ownership space, which is pivotal for following analyses.

²Please note how we need at least one company investing in equity of another company to generate a so-called vertical structure, and hence an ownership network. Otherwise, we would just have a binary relationship between a set of companies and a set of individual shareholders.

³Alternatively, we can think of substituting the equity value with any other firm-level indicator, for example size, productivity, or financial constraints, according to the scope of the analysis. Taking the equity values of companies allows us to better frame the ownership networks as the product of decisions by investors.

⁴In case of lack of fragmentation of shareholding, and companies investing in other companies, we would have vertical hierarchical paths, made of direct paths at full 100% of equity stakes, i.e. each made of one shareholder and one company.

In this framework, it is useful to think of any graph S_G as a collection of ownership paths, running between any of each (direct or indirect) shareholder and a company.

Given the orientation of the ownership relation, any direct shareholder is a (direct) *predecessor* and any receiving company is a (direct) *successor* in ownership. However, there can be more than one predecessor s_{i+m} for each company i , if s_{i+m} is an indirect shareholder whose ownership rights can be extended through m intermediate companies.

Let define the *ownership path* as a finite sequence linking two nodes:

$$P_{s_j i}^{(m)} = (s_j, s_{i+m-1}, \dots, s_i, i) \quad (2)$$

where $s_j \in N$ is a generic predecessor of company i . It can be either an ultimate owner or a corporate shareholder. The *length* of any ownership path is given by m , which is equivalent to the number of predecessors in the path. In case of a *direct shareholder*, the length of the path is $m = 1$.

In fact, alternative ownership paths of variable length may link two nodes in an ownership network. An ownership path $P_{s_j i}^{(m)}$ with $m \neq 0$ can be decomposed in m ownership paths of length 1.

Let define now the *ownership distance* between two nodes as given by the length of the shortest path (or paths) linking the two nodes in an ownership network:

$$d_{ji} = \min_m \left[P_{s_j i}^{(m)} \right] \quad (3)$$

The presence of cycles is a peculiarity of ownership networks, due to cross-holdings and convoluted vertical structures. An ownership network is *cyclic* when it contains at least one ownership path that departs from a (corporate) shareholder and ends back to it. In other words, each company on an *ownership cycle* is at the same time a *predecessor* and a *successor* of itself.

2.3 An illustrative case of ownership network

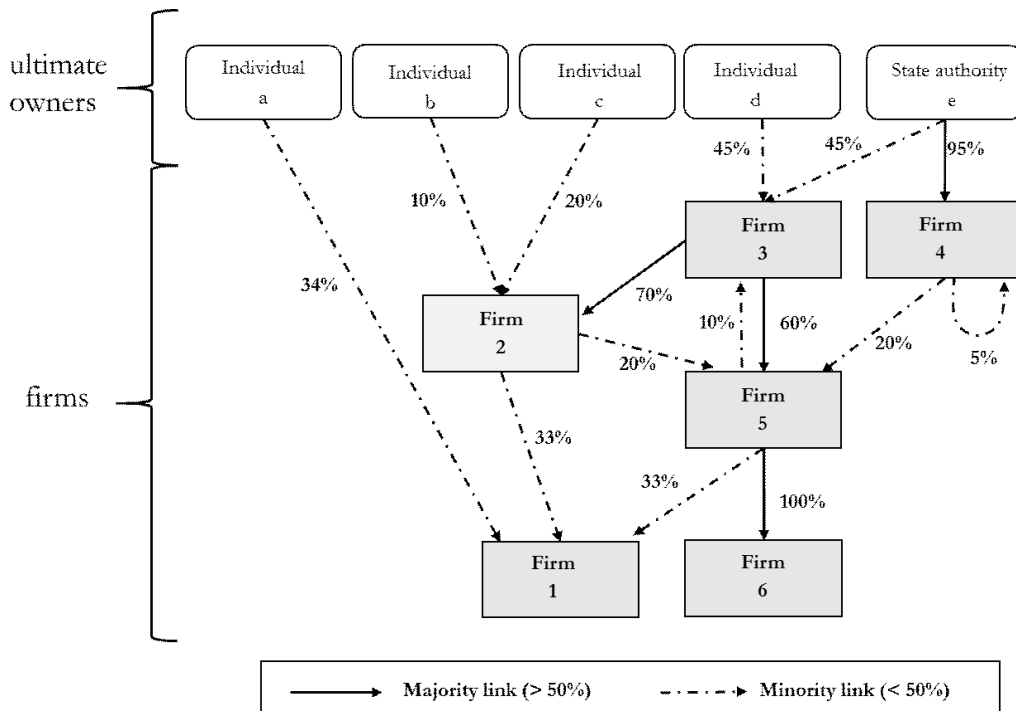
In Figure 3 we report an illustrative case that catches all the hurdles we may encounter in an *ownership network*.

Our illustrative network is made of six companies, four individuals and a state authority (eleven nodes). Overall, fifteen *ownership linkages* are generated, each with an *ownership weight* in percentage. The *ultimate sources (or owners)* of the graphs, the individuals, cannot have any *predecessors* in ownership. We can think of our fictional graph as extracted from a bigger *ownership space*, where other ownership networks are separated. The graph is indeed a *transitive closure*, from and to which no other linkage departs and arrives. That is, no companies and no shareholders are missing from the representation of the network in Figure 3.

Let us consider the ownership relationship between shareholders and companies.

Company [1] has an assembly of shareholders that includes two corporate entities, [2] and [5], and an individual shareholder, [a]. None of its shareholders has outright majority in the company, i.e. each arc has a weight below 50% of voting rights. Company [1] has no successors in ownership and it is at the bottom of any ownership path. Company [2] participates directly in the equity of two different companies, [1] and [5], and it has three distinct shareholders, [b], [c], and [3], among which company [3] has *direct control*. Company [3] has two individual shareholders, [d] and [e], who sit in the assembly together with managers from company [5]. Company [4] is directly controlled by a state authority, [e], but it also holds some of its own shares in portfolio, producing a case of *self-ownership*, which is reported as an arc that both starts and ends in the same company. Company [5] has three corporate entities sitting in the assembly ([2],[3], [4]). Among them, company [3] holds

Figure 3: A stylized ownership network with two corporate boundaries



absolute majority of voting rights. Companies [5] and [3] engage in cross-holdings. Company [6] has only one direct corporate shareholder [5].

Let us consider now a mapping of *ownership paths* in Figure 3.

There are about 54 possible ownership paths connecting any two nodes in the ownership network. For example, company [1] is found at the end of 19 different paths and company [6] at the end of 15 ownership paths. They may have variable length, for example $P_{c1}^{(2)} = (c, 2, 1)$ and $P_{c1}^{(3)} = (c, 2, 5, 1)$ both connect company [1] with individual shareholder [c]. There are four paths with length equal to seven, among them for example $P_{d6}^{(7)} = (d, 3, 2, 5, 3, 2, 5, 6)$. Shortest paths are of direct ownership and have length equal to 1, for example $P_{21}^{(1)} = (2, 1)$. Note how longer paths also contain the ownership cycles, as for example company [2] and [5] appear as predecessors of themselves in $P_{d6}^{(7)}$. A case of self-ownership implies that the company is always considered as a predecessor of itself, as for example in $P_{e6}^{(4)} = (e, 4, 4, 5, 6)$.

All ultimate owners of the network can claim a direct or indirect ownership in company [1]. The same is not true, for example, for company [6], whose ultimate owners are in the subset $\{b, c, d, e\}$.

We can already anticipate here that the *ownership network* represented in Figure 3 contains only one corporate boundaries made by a parent company and its subsidiaries, and a standing-alone firm. It is not always the case. According to our definitions given in Appendix A: i) company [3] is the parent company of [2], [5], [1] and [6] ; ii) company [4] is independent, because it does not belong to any parent company. Albeit independent, company [4] has its own portfolio investment in a subsidiary inside the corporate boundary.

Interestingly, company [6] is controlled by its parent company only indirectly through company [5], i.e. it is a '*subsidiary of a subsidiary*' assuming *transitivity* in control. On the other hand, company [1] is controlled by its parent also after the *consolidation* of voting rights from first-level subsidiaries [2] and [5], which together sum up to an absolute majority of control (66%).

We treat systematically *transitivity* and *consolidation* of voting rights in the next section, as they

are peculiar and sophisticated elements of corporate control structures developed by parent companies to extend the outreach of their corporate boundaries.

3 Corporate boundaries

3.1 Hierarchies of firms

A parent company can coordinate the management decisions of subsidiaries included in its corporate boundary. From this point of view, a corporate boundary can be seen as a set of legally autonomous firms that are under the corporate control of a parent company through shareholding linkages. Reasonably, we can assume that a relationship of corporate control among otherwise independent companies is based on a concentration of voting rights held directly or indirectly by a parent company.⁵

Eventually, corporate boundaries are not one and the same as ownership networks, in the form we introduce in previous Section.2.

On one hand, ownership networks include the case of portfolio investments, when shareholders have an interest in the distribution of dividends but not a long-lasting interest in the management of the company. On the other hand, corporate boundaries exclude *ultimate owners*. In fact, individual investors and public authorities are indeed the original sources of ownership networks and their ultimate economic beneficiaries, but corporate boundaries are based on a separation between ownership and corporate control in modern corporations.

Briefly, we can define a corporate boundary as a hierarchical structure \mathcal{C} , made of a parent and its subsidiaries, extracted from an ownership network \mathcal{G} . A parent and its subsidiaries are in a corporate control relationship, \mathfrak{h} , in the form:

$$\mathfrak{h} : \mathcal{G}(N_{\mathcal{G}}, K_{\mathcal{G}}, V_{\mathcal{G}}, W_{\mathcal{G}}) \rightarrow \mathcal{C}(M_{\mathcal{C}}, L_{\mathcal{C}}, V_{\mathcal{C}}) \quad (4)$$

for every $\mathcal{G} \in \Omega$ and $\mathcal{C} \in \Lambda$, where Λ is the space of corporate boundaries, i.e including all hierarchies of firms.

In other words, the relationship \mathfrak{h} is an application such that from the ownership space, Ω , we extract the corporate control space, Λ . In each ownership network there can be more than one corporate boundaries, not all companies have to be included in a corporate boundary. When a company is not included in the corporate boundary, it is a standing-alone firm with no subsidiaries and no parent.

In this framework, the set of nodes $M_{\mathcal{C}}$ is a subset within the ownership space, such that $M_{\mathcal{C}} \subseteq N_{\mathcal{G}}$. The characteristics of firms falling under hierarchies, $V_{\mathcal{C}}$, are invariant to any control application, although this time the total equity value collected under hierarchies can be lower than the total value distributed in ownership networks, such that $\sum_{i \in \Lambda} v_i \leq \sum_{i \in \Omega} v_i$.

In order to obtain hierarchies of firms from ownership networks, we need an application of corporate control such that:

- i) a unique parent company, $h_o \in N_{\mathcal{C}}$, which controls a set of subsidiaries and it has no *superior*⁶ in corporate control;
- ii) no control cycle in the hierarchy, i.e. no company can indirectly control itself.

Both points define the peculiar characteristics of a hierarchy of firms. The first point assumes that a corporate control relationship must be binary, uniquely attributing subsidiaries to parent companies.

⁵In this we follow international standards. See among others, OECD (2005) and UNCTAD (2009). At this stage we rule out as exceptional the possibility that corporate control is established on contractual terms between a buyer and a supplier, or by using privileged equity shares with control rights in excess of dividend rights. Of course, having information on the latter, it would be easy to extend our model to include these cases.

⁶We adopt the terms *superior* and *inferior* to differentiate the authority structure defined in a hierarchy, from a dominance structure defined in a simple directed ownership network, where we find *predecessors* and *successors*. See also Gilles (2010).

It has to tell us whether a subsidiary is controlled or not by a parent company, *tertium non datur*. It implicitly states that any simple portfolio investment is ruled out from a corporate boundary.

The second point assumes that we have to find a way to untie bundles of ownership cycles (cross-holdings, cross ownership, and self-ownership) in order to understand which is eventually the parent company on top of the hierarchy.

Eventually, a hierarchy is an *unweighted* graph, where no percentage is attached to *control linkages*, L_C . We will better define the nature of control linkages in following paragraphs, after an introduction to the corporate control application.

3.2 Corporate control and interlocking assemblies

Corporate control is based on a concentration of voting rights. The simplest case is when a unique shareholder has absolute majority in a company, but when ownership is fragmented a coalition of shareholders is needed. In fact, building coalitions in ownership networks is complicated by the presence of interlocking assemblies of shareholders.

Assemblies can be interlocking because a shareholder can have a right to sit at more than one assembly, and because a corporate shareholder has to rely on the decision taken by its own shareholders before taking a decision in the assembly of another company. Eventually, we may find cascades of votes in ownership networks, which are more sophisticated when cross-holdings and long vertical structures have to be taken into account.

In this framework, we can consider any hierarchy of firms, \mathcal{C} , as a collection of control paths connecting nodes. This allows us to express a relationship of control \mathfrak{h}^+ in the form:

$$\mathfrak{h}^+(h_0) = \left\{ i \in M_C \mid \exists P_{h_0 i}^{(m)} = \{h_0, h_{i+m-1}, \dots, h_i, i\} \text{ in } \mathcal{C} \right\} \quad (5)$$

where h_0 is the parent company and the control path has a generic length m . Any generic element h_{i+m-n} , with $n > m$, is an intermediate controller of company i .

The sign of the mapping, \mathfrak{h}^+ , defines the orientation of the relation from *superior* nodes to *inferior* nodes, as we assume that headquarters by parent company is where coordination of management starts. That is, we assume that once a decision is made at the headquarters level, this is passed along *control paths*. There can be more than one control path running from a parent and each of its subsidiaries, with different lengths.

In this case, we define the *hierarchical distance* of a subsidiary from the parent company as the minimum length of control paths, in the form:

$$d_{h_0 i} = \min_m \left[P_{h_0 i}^{(m)} \right] \quad (6)$$

In other words, a corporate boundary has shareholders' assemblies that are interlocking and possibly coordinate with headquarters, but they are ordered according to a relationship that we can consider of partial order⁷, such that $h_0 \succeq h_{i+m-1} \succeq \dots \succeq i$. In fact, there can be more than one control path with the same distance from the headquarters. Also, if two companies exchange majority cross-holdings, they can be considered one the superior of the other. It would be clear from following analyses how a relationship of partial order catches the complexity of some corporate governance arrangements, better than a strict order relationship.

⁷According to properties of partial order: i) any company can be a superior of itself (reflexivity: $h_m \succeq h_m$); ii) if a company controls another and the latter control the first, they are on the same hierarchical level (antisymmetry: if $h_m \succeq h_{m+1}$, and $h_{m+1} \succeq h_m$, then $h_{m+1} \sim h_m$); iii) if a company controls another, and the latter in turn has a subsidiary, then the company on top control both (transitivity: if $h_m \succeq h_{m+1}$, and $h_{m+1} \succeq h_{m+2}$, then $h_m \succeq h_{m+2}$).

3.2.1 Transitivity of control and consolidation of voting rights

We assume that a parent company can extend its outreach on an ownership network in two ways:

- i) by transitivity of control, establishing vertical structures, i.e. allowing subsidiaries to control other subsidiaries;
- ii) by consolidation of voting rights, summing up the shares held by its subsidiaries in other companies to reach majority of voting rights in the latter.

However, the extent of a corporate boundary can be unknown to the analyst. In this case, one needs a full picture of the interlocking shareholding among companies, to detect and assign corporate boundaries. It was the case of Figure 3, where only a full mapping allowed us to track the extension of a corporate boundary from an ownership network. In Appendix Figure B1 we report an application of the following methodology to the fictional network of Figure 3.

We proceed recursively, in the spirit of what proposed by Chapelle and Szafarz (2007)⁸, bottom-up from ownership of single companies.

We start considering an oriented square matrix with all direct shareholding activity represented in a network \mathcal{G} , as the one in eq. (1), where each entry corresponds to the equity stake $w_{s_i i}$, if s_i is a direct shareholder and 0 otherwise.

Hence, we transform the original ownership matrix in a matrix $A^{(1)}$, whose single element is:

$$a_{s_i i}^{(1)} = \begin{cases} 1 & \text{if } s_i : w_{s_i i} > 0.5 \\ 0 & \text{if } \exists s_k \neq s_i : w_{s_k i} \leq 0.5 \\ w_{s_i i} & \text{otherwise} \end{cases} \quad (7)$$

Briefly, matrix $A^{(1)}$ identifies all the *immediate controllers* (if any) among direct shareholders who reach absolute majority in a company i , $w_{s_i i} > 0.5$.

If an immediate controller is identified, any minority shareholders s_k does not have a chance to participate to corporate control and we simplify dropping their linkages. Otherwise, if no absolute majority is detected, we keep the linkages of all the shareholders in a company i as they are.

Transitivity of control at upper levels

It is possible that a (corporate) direct controller is in turn controlled by another company. In this case, we consider shareholding at an upper level with respect to company i , i.e. looking for shareholders of direct (corporate) controllers'. We build a new matrix, $T^{(1)}$, whose element is:

$$t_{s_j i}^{(2)} = \begin{cases} 1 & \text{if } \exists P_{s_j i}^{(2)} = (s_j, s_i, i), \text{ where } w_{s_j s_i} > 0.5 \\ 0 & \text{if } \exists P_{s_k i}^{(2)} = (s_k, s_i, i), \text{ where } s_k \neq s_j \text{ and } w_{s_k s_i} > 0.5 \\ a_{s_i i}^0 & \text{otherwise} \end{cases} \quad (8)$$

where s_j and s_k are to be found among all 'direct shareholders of direct shareholders'. All the generic paths $P_{s_* i}^{(2)}$ of length two, reaching the company and starting from shareholders of direct controllers, s_j and s_k have to be taken into account.

Consolidation of shares at upper levels

It is also possible that on upper levels we can consolidate to majority shares that are dispersed among different (corporate) shareholders of company i , which actually belong to the same controller, as in the case of Figure 3. We build a matrix $A^{(2)}$, such that each of its element is:

⁸Chapelle and Szafarz (2007) apply an algorithm similar to ours on already defined corporate boundaries, to assess the power of companies inside the boundary. Our scope is different. We want to detect the extent of corporate boundaries when these are unknown, given the information on raw ownership. In this case, we apply Banzhaf index after (not before) consolidation of control, hence using power indices in a more conservative way.

$$a_{s_j i}^{(2)} = \begin{cases} 1 & \text{if } t_{s_j i}^{(1)} + \sum_{k: t_{s_j s_q}^{(1)}=1} t_{s_q i}^{(1)} > 0.5 \\ 0 & \text{if } \exists s_l \neq i : t_{s_j i}^{(1)} + \sum_{l: t_{s_j s_l}^{(1)}=1} t_{s_l i}^{(1)} > 0.5 \\ t_{s_j i}^{(1)} & \text{otherwise} \end{cases} \quad (9)$$

where s_q and s_l are to be found among predecessor of i on paths of length 1, i.e. its direct shareholders.

From now on we iterate consecutively, looking always at a generic upper level m of shareholding to detect transitivity of control in a matrix $T^{(m)}$. Then we look downstream at the generic level $(m-1)$, to detect consolidation of control in a generic matrix $A^{(m-1)}$. We finally come up to the point that $A^{(n)} = A^{(n+1)}$, which implies that there is no ownership path longer than n that allows to detect either transitivity or consolidation of control.

At the end of the process, we know the extent of corporate boundaries, \mathcal{C} 's, which are identified by all control paths $P_{h_o i}^{(m)} = \{h_o, h_{i+m-1}, \dots, h_i, i\}$ with a (corporate) controller on top, h_o . We discard ultimate owners on top of parent companies, because they do not belong to the corporate boundary.

3.3 Coalitions of shareholders and control probability

So far, we have left out from corporate boundaries minority stakes, once assuming they represent portfolio investments, i.e. the shareholder does not participate to the control of a company because there is a well established control path starting from parent companies.

Yet we may find still companies, whose ownership is fragmented, where (individual or corporate) shareholders can form coalitions and pool together their voting rights to have an impact on the management. Indeed, this is the case of quoted companies, where one or few shareholders can gain actual control with just a minority stake, because all the rest of share capital is floating on the market and fragmented among an ocean of small shareholders (see for example Nenova, 2003, and Rajan and Zingales, 1998a). These cases are not encompassed by transitivity and consolidation of control as introduced in the previous section.⁹

For these cases, we can extend the notion of corporate control adopting a probabilistic measure comprised in a range $[0, 1]$, whose limit value 1 is the case of a unique majority shareholder.¹⁰ Thereafter, at the end of this section we discuss peculiar cases of coalitions of shareholders, which can shape real-world cases of a wedge between ownership and control in companies.

3.3.1 The Banzhaf (1965) Index

We have an assembly of direct shareholders, S_i , of company i that adopts a voting rule to reach consensus on each decision. The preferred majority rule is binary, because shareholders are asked to vote 'yes' or 'no' on single management proposals. We define the outcome of each possible vote in company i as $r(B_i) \in \{0, 1\}$, where $B_i \subseteq S_i$ is a winning coalition, and the 'yes' and 'no' are respectively proxied by 1 and 0. Conventionally, we assume that a coalition wins when its proposal

⁹Until now we followed international standards for concentration of voting rights above 50% by a unique shareholder, with or without consolidation and transitivity. For an example of a more flexible definition in international law, we may refer to the Article 3(2) of the EU Merger Regulation, where control is defined as the possibility of exercising decisive influence on an undertaking, which can be acquired through purchase of securities or assets or by rights, contracts, or any other means (see http://ec.europa.eu/competition/mergers/legislation/merger_compilation.pdf)

¹⁰Please note how the use of a Banzhaf Index in previous works (Levy and Szafarz, 2017; Levy, 2011; Crama and Leruth, 2007) is applied to already defined corporate structures in order to assess the distribution of control power within the structure. Here we adopt a different perspective, as our scope is to identify corporate structures that are embedded in ownership networks. In our case the Banzhaf Index is adopted after (and not before) consolidation and transitivity in majority of voting rights, to attribute residually to a corporate boundary the firms that are otherwise left outside any corporate control network.

is approved ('yes') by the absolute majority of voting rights ($> 50\%$). The generic coalition B_i is winning if:

$$\sum_{s_i \in B_i} w_{s_i i} \in (0.5, 1] \quad (10)$$

The outcome of voting session eventually depends on the distribution of weights, $w_{s_i i} \in W_i$, among shareholders. For sake of completeness, we include among possible coalitions also the empty set, when no agreement is reached ($B_i = \{\emptyset\}$), and the *grand coalition* ($B_i = S_i$), when all agree on a proposal.

We are able to measure the control probability of each shareholder, $\pi(s_i)$, after the computation of a Banzhaf (1965) index¹¹ in the form:

$$\pi(s_i) = \frac{1}{2^{|S_i|-1}} \sum_{B_i \subseteq S_i \setminus \{s_i\}} r(B_i \cup \{s_i\}) - r(B_i) \quad (11)$$

where $|S_i|$ is the cardinality of the corporate assembly S_i . The control probability, $\pi(s_i)$, falls in a range $[0, 1]$. It is the probability of a shareholder s_i to exert the decisive vote, once considering all the possible winning coalitions with her, $B_i \cup \{s_i\}$, and without her, B_i .

A shareholder has probability 1 of controlling a company if she has alone above 50% of equity stakes. In this case, any other shareholder would have a control probability equal to 0. When no shareholder has absolute majority, different distributions of control power can be detected between 0 and 1, which are not linear in the distribution of equity shares.

3.3.2 Dominant stakes

As a consequence of eq. 11, when companies have fragmented ownership and nobody has outright absolute majority, the power of one shareholder to influence management depends also on the distribution of capital shares among the others. The more fragmented is the ownership among other shareholders, the higher the control probability by one shareholder. Take for example the case of companies that are quoted at the stock exchange, where one can obtain control with a threshold well below 50% of voting rights. This is possible because a floating of stakes exists that can be distributed among a mass of investors that is not able to unite and form a coalition. In general, we can say that a shareholder has an influential stake when its control probability is high enough, even when a unique controller has not been identified.

Then, what is the threshold of voting rights after which a stake in a company is influential? In a probabilistic framework like the one given by eq. 11 there is no single answer.

In Appendix Figure B2, we report a correlation between equity shares held by a shareholder and her Banzhaf index, as derived from an application of eq. 11 to the ownership data we are going to use throughout the rest of the paper. There we find that there is no linear association between the amount of voting rights (on the x-axis) and the probability of controlling a company (on the y-axis), as soon as no shareholder collects alone $> 50\%$ of shares. A variety of situations can be found, depending also on the number of shareholders sitting at a firm assembly.

Nonetheless, we can reasonably assume that a shareholder is *dominant* when her Banzhaf Index is at least higher than 0.5, i.e. her voting rights are determinant in at least half of the possible coalitions in an assembly of shareholders.

¹¹This index is also called the Penrose-Banzhaf-Coleman power index. For previous use of the theory of cooperative games in detecting shareholding and corporate structures, see the review by Crama and Leruth (2013). Two main power indices are originally designed for horizontal voting models. The alternative is the Shapley value (1953) index, which calculates the fraction of the possible voting *sequences*, in which a shareholder is decisive for the approval. Our preference falls on the Banzhaf (1965) index as it is valid whatever the voting sequence. See also Dubey and Shapley (1979) for an introduction to its mathematical properties.

We consider the notion of dominant stake as just complementary to transitivity and consolidation of voting rights in defining the extension of corporate control networks. We illustrate hereby some peculiar situations that intervene to define better some exceptional cases we can find in real-world corporate hierarchies.

A) A dominant stake can be held by a parent company and/or one of its subsidiaries on an otherwise independent company.

By transitivity of control, we can assume that the parent of the subsidiary finally exerts its power also on the company in which it and/or its subsidiary have a dominant stake. In the latter case, a further coordination effort is required between the parent and the subsidiary, which makes control more disputable by other shareholders.

B) A dominant stake can be held by a parent company on another parent company.

Although rare, it is possible. This is the case of the Toyota Group described from introduction in Figure 2, where eight parent companies were spotted. They have a complex cross-holdings structure in each other. For example, Toyota Motor Corporation has a stake between 20% and 25% in each of the other seven parent companies, while two of the latter (Toyota Industries Corporation and Denso Corporation) have respectively a stake of 6.63% and 2.56% in Toyota Motor Corporation. Considering fragmentation and cross-holdings, for example, Denso Corp is dominated in corporate control by a coalition made by Toyota Motor Corporation and Toyota Industries Corporation, which respectively control 22.25% and 10.32%, i.e. not an absolute majority of voting rights

C) A dominant stake can be held by a financial company on an industrial company.

Financial companies may professionally invest in the equity of other companies, ending up with an important amount of stakes in many companies, although they may be not committed in coordinating the management decisions of these firms. Indeed, according to the level of commitment and depending on the regulatory environment, a financial company can act as an active or a passive investor. This is for example the case of pension funds and mutual funds, which may become more active in case of extraordinary decisions in the life of companies after exerting pressure on managers of the company.

D) A dominant stake can be held by an ultimate owner on a parent company.

In this case, the ultimate owner with the highest stake is the major beneficiary of the activity of the parent company and its subsidiaries, but she does not participate to the definition of the corporate boundary by virtue of separation between ownership and control. She will participate proportionally to the distribution of dividends, whereas her control probability in the company defines her power with respect to the other ultimate owners. We can find under this category for example family business groups, when the equity is fragmented among more than one member of the family. Take the case of Tata Motors Limited originated in India, which is generally considered the parent company of the group, although members of the Tata family in India own the conglomerate group with just a direct share equal to 31.06% in the portfolio of a family holding, the Tata Sons company.¹²

E) A dominant stake can be held by an ultimate owner on an independent company.

The company is still independent, as it does not participate to any corporate control network, because it does not control any subsidiary. As in the previous case, any ultimate owner is the beneficiary of the activity of the independent company in proportion to the capital shares, but she will have a higher power the higher her Banzhaf Index. This is the case of some partnerships, when individual partners decide to share the risk of the business venture among them.

¹²The rest of the equity is distributed among professional financial investors as Citibank, which is actually the second shareholder with a share of 18.37%. Members of the Tata family ultimately endow two main trusts created by the family founder with 66 percent of the equity capital of Tata Sons Company for philanthropic actions, the Sir Dorabji Tata Trust and the Sir Ratan Tata Trust. We can easily consider the whole shareholding activity detected on top of the Tata Motors Limited as alien to the productive corporate boundary of the Indian conglomerate.

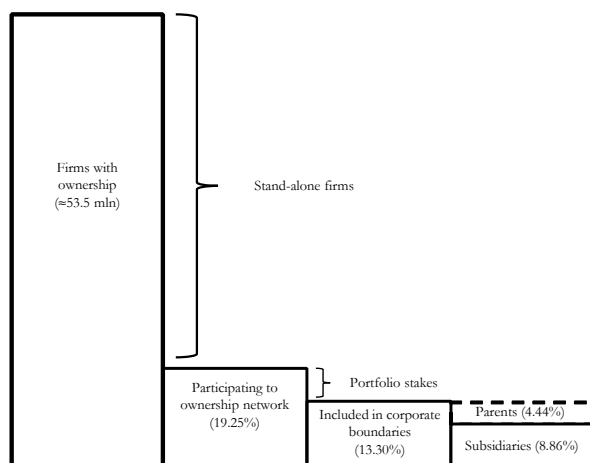
4 Data and preliminary evidence

4.1 Data on ownership networks

We source original information on global ownership from Orbis, a database compiled by Bureau Van Dijk, a private publisher of business information, which collects shareholding information on more than 53.5 million companies operating in 208 countries, which are active at the end of the year 2015. The ownership information is originally retrieved from a variety of national and international information providers (national registries, regulatory bodies), as well as from company annual reports, company websites and specialized press.

In Figure 4 we start from full coverage of the ownership sample and separate stand-alone companies, i.e. companies that are not participated by any corporate shareholder and that are not corporate shareholders of any other company in our data. These are companies that do not participate to ownership networks. Nested in Figure 4, we end up with a subset of 10,294,391 firms¹³, which are instead embedded in ownership networks, because they participate or are participated in capital shares.

Figure 4: Relevance of corporate control networks on total ownership - number of firms



Among them, 3,184,147 firms engage in portfolio investment operations, because they are not controlled by a corporate shareholder and do not reportedly control any subsidiary¹⁴. Their corporate shareholders have only minority stakes, and/or they have only minority stakes in the share capital of other companies.

Therefore, after the application of the algorithm proposed in Section 3, we detect corporate control relations among 2,369,892 parent companies and 4,740,352 subsidiaries.¹⁵

In Table 1, we provide a geographic coverage of our dataset made of parents and subsidiaries. In Appendix Table B1, we also report the geographic coverage and firm size details of the more complete dataset, including portfolio investment. Throughout our analyses, we follow the procedure of Section 3 and we adopt the definitions provided in Appendix A.

¹³The majority of companies not participating to ownership networks (62%) in 2015 is made of single traders/proprietorships, non-profit organizations and other unlimited liability companies. Limited liability companies that do not participate in other companies are also excluded (38%). However, unlimited liability companies can be shareholders of limited liability companies and hence be present in the dataset.

¹⁴In Appendix A we provide details on the definitions for firms and shareholders (subsidiary, affiliate, parent company, ultimate owners) in line with international standards for national accounts.

¹⁵At this stage, we assume that one parent company uniquely identifies also one corporate control network. Please, see section 3.2.2 for a discussion of peculiar cases, in particular when more than one parent company is present in the corporate boundary, as in the case of the Toyota Group reported in introduction.

Table 1: Companies participating to corporate control networks. Geographic coverage and multinational status.

Hosting economy	Parent companies				Subsidiaries			
	All	%	Multinational	%	All	%	Foreign	%
European Union	600,829	25.35%	111,522	55.41%	1,625,508	34.29%	387,006	51.44%
<i>of which:</i>								
<i>Germany</i>	58,969	2.49%	11,261	5.59%	196,426	4.14%	64,405	8.56%
<i>France</i>	31,389	1.32%	6,650	3.30%	96,749	2.04%	26,631	3.54%
<i>United Kingdom</i>	58,138	2.45%	12,361	6.14%	285,286	6.02%	62,818	8.35%
<i>Italy</i>	40,555	1.71%	8,680	4.31%	88,091	1.86%	19,357	2.57%
<i>Spain</i>	34,600	1.46%	5,530	2.75%	103,454	2.18%	21,823	2.90%
United States	1,435,218	60.56%	22,511	11.18%	2,138,025	45.10%	63,220	8.40%
Russia	29,741	1.25%	974	0.48%	110,232	2.33%	50,541	6.72%
Other Europe	36,073	1.52%	14,089	7.00%	84,045	1.77%	22,441	2.98%
Asia	105,449	4.45%	19,142	9.51%	316,014	6.67%	99,624	13.24%
<i>of which:</i>								
Japan	32,526	1.37%	3,259	1.62%	82,316	1.74%	5,214	0.69%
China	11,048	0.47%	2,995	1.49%	83,311	1.76%	35,983	4.78%
India	3,516	0.15%	1,501	0.75%	14,971	0.32%	6,775	0.90%
Africa	5,102	0.22%	4,169	2.07%	30,346	0.64%	17,088	2.27%
Latin America	30,058	1.27%	18,247	9.07%	83,227	1.76%	51,693	6.87%
<i>of which:</i>								
Brazil	2,905	0.12%	342	0.17%	15,960	0.34%	9,443	1.26%
Argentina	600	0.03%	126	0.06%	3,907	0.08%	3,059	0.41%
Mexico	1,356	0.06%	322	0.16%	12,728	0.27%	8,132	1.08%
the Caribbean countries	13,023	0.55%	12,662	6.29%	21,684	0.46%	18,049	2.40%
Australia	58,788	2.48%	2,771	1.38%	136,189	2.87%	14,750	1.96%
Rest of the world	68,634	2.90%	7,847	3.90%	216,766	4.57%	45,992	6.11%
TOTAL	2,369,892	100.00%	201,272	100.00%	4,740,352	100.00%	752,355	100.00%

In fact, firms of any size ¹⁶ are included. Overall, small firms represent 86.6% of the whole dataset and 40.6% of small firms are present in US. Global ownership spans across 208 countries in our dataset but, as expected, the most represented economies are the European Union and the United States, with respectively a share of 38.4% and 36.2% of the total dataset. We have 7.8% of Asian companies and among them Japan and China add to 40% in their continent. Russia has a share of 5.6% of the entire dataset and 13.5% of the total in Europe. The less active geographic areas are Africa and Latin America. Small, medium and large companies are well represented in ownership networks.

Numbers by country/area slightly change in Table 1, after we exclude portfolio operations. United States host by far the highest numbers of both parent companies (60.56%) and subsidiaries (45.10%). The United Kingdom and Germany are respectively the second and third country with the highest numbers of parents (2.49% and 2.45%) and subsidiaries (4.14% and 6.02%). Asia is the third area most active, hosting however only 4.45% of parents and 6.67% of subsidiaries in the world.

In Table 1, we further detect the multinational nature of corporate control networks. We consider a parent company as leading a multinational enterprise (MNE) if it controls at least a foreign subsidiary that is located in a country different from the country origin of the parent. As expected, in the European Union we detect a higher share of multinational enterprises (55.41%) than in US for two reasons. First, there is a high level of integration at the European continental level, because companies locate activities intra-EU in neighboring countries. On this respect, companies inside US investing from one state to another are certainly not considered as multinational enterprises. Second, MNEs in Europe are usually smaller in size than in US, and consequently more numerous in Table 1. In fact, in US the average number of subsidiaries controlled by an MNE is about 12, whereas in the European

¹⁶A combination of criteria is adopted by original sources to classify a company by size: A) Large or very large companies register more than 10 mln EUR revenues, or more than 20 million EUR total assets, or more than 150 employees, or over 0.5 mln EUR capitalization, or they are listed at a stock exchange; B) Medium-sized companies register more than 1 mln EUR revenues, or more than 2 million EUR total assets, or more than 15 employees, or over 50 thousand EUR capitalization; C) Small companies includes the residual, i.e. they are not in the medium or in the large and very large categories.

Union is only about 7. Nonetheless, the United States host the highest number of affiliates, of which a great majority is actually controlled by domestic parent companies.

Around the world, we detect 752,355 foreign subsidiaries, which is about 15.8% of the total. More than half of them are located in the European Union. Asia is the second continent with the highest number of foreign affiliates (99,624), 36% of which is hosted in China. Among Latin American countries, a noteworthy 2.4% of world share of foreign affiliates is attracted by Caribbean countries, especially in British Virgin Islands, Bahamas and Curacao, well in excess of their economic weight. Of these, many are peculiar cases of blurring nationality and round-tripping investment, as discussed in following Section 4.5.

4.2 Ownership fragmentation and cross-holdings

We provide here some descriptive statistics on the fragmentation of global ownership and on the relevance of the phenomenon of cross-holdings among companies.

Basic network indicators based on ownership linkages are reported in 2. In fact, ownership linkages are the elementary observation units shaping the entire ownership space.

Each linkage can be represented as an oriented edge between a shareholder and a company. In the first column, the *ownership in-degree* is essentially the number of shareholders by firm. In the third column, the *ownership out-degree* is the number of capital shares held by shareholder. Columns 2 and 4, respectively, represent the *ownership in-strength* and *out-strength*, i.e. the weights of stakes (as percentages) in each participated firm and as held by each shareholder. Please note how the category of shareholders includes individual, corporate, and state shareholders.

Table 2: Ownership distribution: in-degree, in-strength, out-degree and out-strength

	N. Shareholders by owned firm (in-degree)	Capital share (%) by owned firm (in-strength)	N. Stakes by shareholder (out-degree)	Capital share (%) by shareholder (out-strength)
mean	1.75	66.46	1.56	43.12
5th percentile	1	20.00	1	0.32
10th percentile	1	33.33	1	1.13
25th percentile	1	50.01	1	10.00
median	1	50.01	1	50.01
75th percentile	2	100.00	1	52.95
90th percentile	3	100.00	2	100.00
95th percentile	4	100.00	3	100.00
99th percentile	10	100.00	8	100.00
min	1	0.01	1	0.40
max	2,050	100.00	10,618	100.00
st. dev.	4.59	28.77	12.12	34.06
skewness	73.06	0.01	389.07	0.40
N. firms/shareholders	10,294,391	10,294,391	8,197,736	8,197,736
N. ownership linkages	13,251,862	13,251,862	13,251,862	13,251,862

In Table 2, the median firm has only one shareholder and the median shareholder owns stakes in just one company. However, both the distributions of *ownership in-degree* and *out-degree* show high skewness. After the 75th and 90th percentiles, respectively, there is a rapidly increasing fragmentation of shareholding activity.¹⁷

¹⁷The firm with the highest number of shareholders in 2015 is Rizo Consorzio Servizi SCARL, a limited liability Italian consortium mainly involved in transportation and storage activities, with 2,050 shareholders. In the same period, the company with the highest number of stakes in its portfolio is Dimensional Fund Advisors, an American financial company that manages assets for mutual funds. It owns stakes in 10,618 different companies.

Interestingly, in the right tails of these distributions we usually find bigger companies. Bigger companies have on average more shareholders, because the presence of a second shareholder correlates with the probability that a company is three times bigger than one with a unique shareholder. Also, bigger companies have more capital shares, because the possession of a second stake in portfolio correlates with the probability that a corporate shareholder is four times bigger than one with just one stake.

On the other hand, if we look at the distribution of *ownership in-strength* in Table 2, we find that both the median and the average companies are owned with an absolute majority (50.01% and 66.46%, respectively). *Ownership out-strength* is relatively more fragmented than *in-strength*, because an average (individual or corporate) shareholder has a 43.12% stake in a company. Financial companies are responsible for keeping lower the average capital share, because they professionally invest small stakes in equity of different companies to diversify their portfolio risk.

Nonetheless, we find that about 1.9 million of firms in our data do not have a unique shareholder able to exert absolute majority (>50%) of voting rights. In these cases, either a coalition of shareholders as in Section 3.3 or a consolidation of voting rights in the form introduced in Section 3.2.1 are needed to establish corporate control.

Interestingly, we can also assess the phenomenon of cross holdings, when two companies hold capital shares in one another. We find that 16,794 companies in 103 different countries are involved in cross-holding activities. In Appendix Table B2, we show descriptives for the top 20 countries where cross-holdings is detected. Unexpectedly, German is by far the first with 4,956 companies, followed by US at distance with 1,556 companies. Japan scores only third in the ranking, although the country is usually considered as the principal case study for cross-holdings¹⁸ (Bohren and Michalsen, 1994). However, all major EU members find a place in the list represented in Appendix Table B1. Among them, German companies present the highest average cross-holding share (93%), whereas Japanese companies have an average cross-holding share of just 14%, the latter more in line with the goal of establishing blocking minorities and/or cementing business alliances, rather than for exerting full corporate control.

4.3 Corporate control and economic entrenchment

After transitivity and consolidation of voting rights¹⁹, we obtain corporate control networks whose distributions is plotted in Figure 5. On the left panel we document all cases, whereas on the right panel we report only MNEs.

We measure the size of control networks as the (log of the) number of its affiliates²⁰, on the x-axis, and we report the (log of the) count of how many control networks have that size, on the y-axis. If we assume that Figure 5 underlies a negative binomial distribution, we estimate that (over)dispersion is equal to 1.58 for all corporate networks in the first panel, and it is 2.89 for multinationals in the second panel.

That is, the majority of these objects has very simple structures, whereas only a few have sophisticated structures. In fact, on the left panel of Figure 6, we count 74.91% of cases made of only two firms, when one parent company owns a single subsidiary (59.29% in the case of multinationals), whereas only 0.08% have more than 100 foreign or domestic subsidiaries. These are 1,934 groups of firms, which however coordinate the majority of economic activity on a global scale. Indeed, they account for more than half of the turnover produced in the total sample²¹, as reported in the right panel of Figure 6.

¹⁸In the same year 2015, the new Japanese governance came into force to make corporate governance of Japanese firms more transparent. It probably acted already as a deterrent to reduce the phenomenon of cross-holdings.

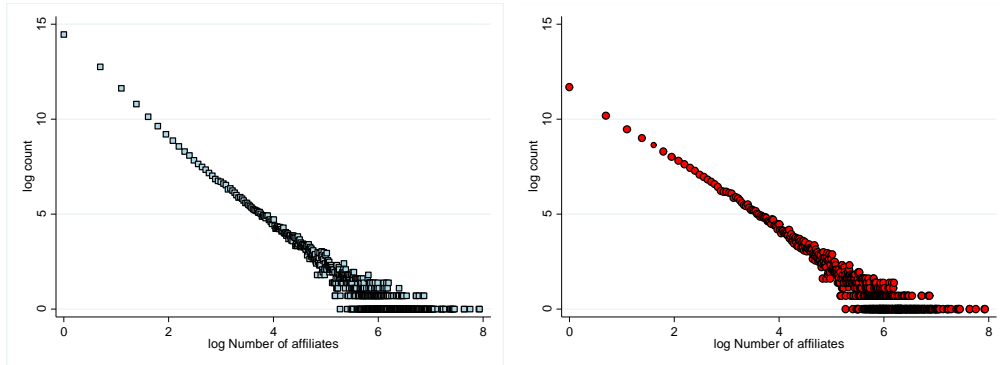
¹⁹In following analyses, we propose results excluding affiliates where parent companies have dominant stakes, due to the possibility of double counting, when control probability is equal across more than one shareholder. However, main findings do not substantially change. Graphs and tables including cases of dominant stakes are available upon request.

²⁰Assuming that each corporate control network is identified by a unique parent company. See Section 3.3. for a discussion of peculiar cases of multi-parent corporate groups. See the case of Microsoft Corp reported in introductory Figure 1.

²¹Turnover is present in less than half of the sample of subsidiaries. We may assume that smaller firms, in smaller networks, can affect our findings on concentration. Nonetheless, our descriptive finding is robust to the exclusion of all

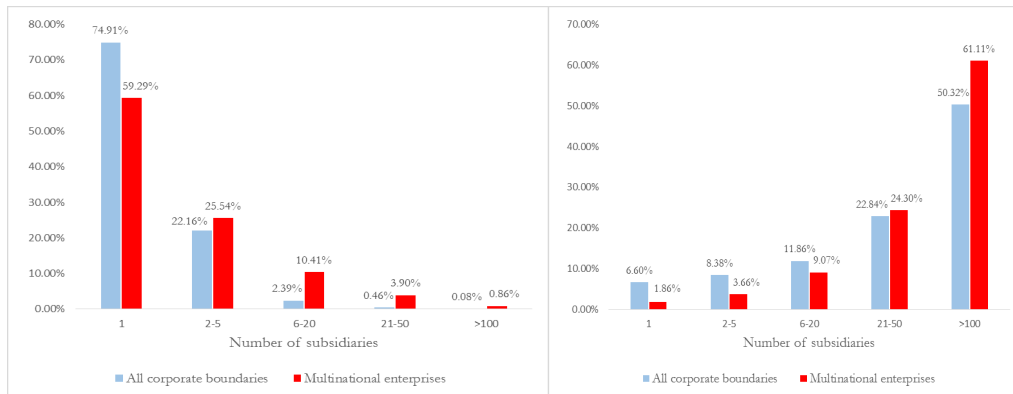
In general, multinational activity is even more concentrated, since multinational groups are on average bigger in terms of number of subsidiaries, while less than 1% of them explains more than 60% of multinational sales.

Figure 5: Distributions by size of corporate networks



a) All corporate networks (2,369,892). We estimate dispersion $(1+\alpha)$ parameter 1.58 assuming a negative binomial distribution. b) Multinational corporate networks (201,272). We estimate dispersion $(1+\alpha)$ parameter 2.89 assuming a negative binomial distribution.

Figure 6: Distribution of size and sales in corporate networks



a) number of corporate networks by size category (as numbers of subsidiaries) b) value generated by size category (as numbers of subsidiaries)

corporate networks where information on turnover is not complete. Still, we can confirm that less than 1% of these objects concentrate around 50% of global sales.

4.4 Hierarchical distance

The main advantage we have when adopting a network perspective in corporate control is the possibility to observe paths and distances.

A parent company can be connected in many ways with a subsidiary, through one or more proprietary linkages at variable distance from the headquarters. More than one subsidiary can be present on a control path. At the end, each *middlemen* subsidiary on a control path may have different characteristics, which can explain why the parent company decides to go right through it before controlling another subsidiary.

Paths can run across different countries, with different characteristics and varying institutional frictions, therefore the classification by nationality of a firm, i.e. the identification of its origin country, may be difficult.

We rely on our framework, and specifically on eq. (6), to attribute a control distance between a parent and each of its subsidiaries, defined as the length of the shortest path through the ownership network. Our definition allows to assume that companies want to minimize the communication costs entailed in passing management decisions downstream from the headquarters. Ideally, in a coordinated sequence, a decision starts from the parent company and is enforced first by its immediate subsidiary, which in turn pass it to subsidiaries of subsidiaries. If there is any crossroad on the ownership network, the parent company will prefer to take a shortcut and reduce frictions.

In Table 3, we plot hierarchical distances of subsidiaries as an important measure to understand the organization of corporate control in a hierarchy of firms.

An overwhelming majority of companies (78.6%) are located at distance one from the parent company. However, *middlemen* subsidiaries and indirect control paths are detected in a non-negligible 21.5% of cases. In the case of foreign subsidiaries the figure of indirectly controlled affiliates goes up until 53.7% of the total. For about 2,000 subsidiaries, we can find more than ten *middlemen* subsidiaries linking a final subsidiary with its headquarters.²²

Table 3: Hierarchical distance of domestic and foreign subsidiaries from parent company

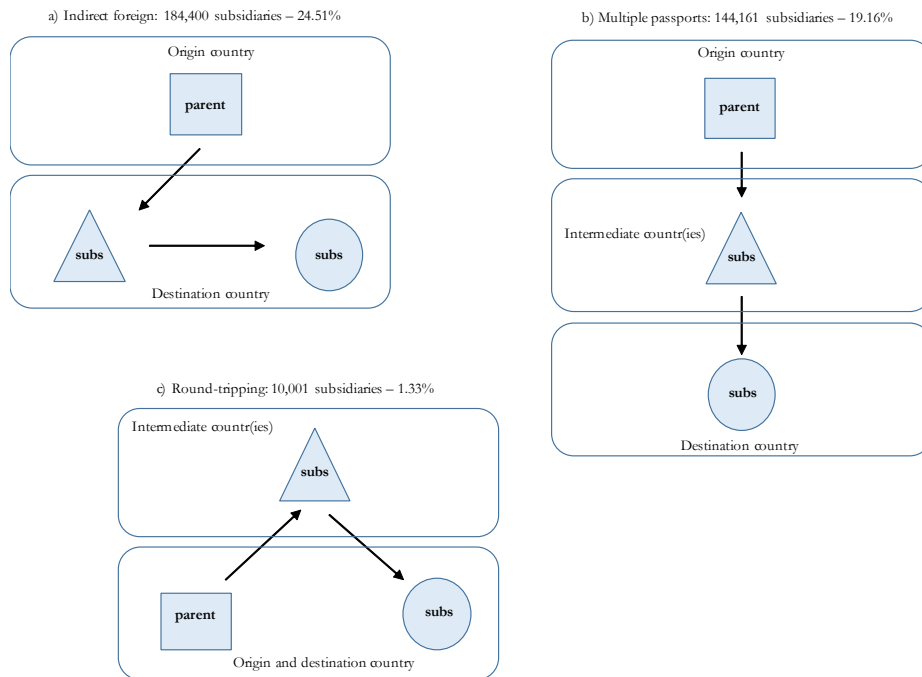
Hierarchical distance	Domestic subsidiaries		Foreign subsidiaries		All companies	
		%		%		%
1	3,223,182	84.991%	347,959	46.249%	3,571,141	78.578%
2	376,511	9.928%	186,062	24.731%	562,573	12.379%
3	123,021	3.244%	104,063	13.832%	227,084	4.997%
4	43,079	1.136%	55,414	7.365%	98,493	2.167%
5	15,354	0.405%	28,135	3.740%	43,489	0.957%
6	5,934	0.156%	14,182	1.885%	20,116	0.443%
7	2,518	0.066%	8,132	1.081%	10,650	0.234%
8	1,321	0.035%	3,765	0.500%	5,086	0.112%
9	600	0.016%	2,104	0.280%	2,704	0.059%
10	268	0.007%	987	0.131%	1,255	0.028%
> 10	580	0.015%	1552	0.206%	2132	0.047%
Total	3,792,368	100.000%	752,355	100.000%	4,544,723	100.000%

²²As an example of extremely long control path, we report the round-tripping case of Europarks U.K. Limited, which is located at hierarchical distance twenty from its parent company at the end of 2015. In fact, its control path crosses the border twice, in Luxembourg and in UK, where we also found its parent company Bottom-up. In between, we have: Europarks U.K Limited - Europarks Limited - National Car Parks Limited - National Parking Corporation Limited - NCP Holdings Limited - National Car Parks Group Limited - Pointspec Limited - Statusaward Limited - Trendcycle Limited - Primepanel Limited - PIHL (2003) Limited - Oval (2041) Limited - Oval (2042) Limited - Parking International Holdings Limited - Parking Holdings Limited - MEIF II CP Holdings 3 Limited - MEIF II CP Holdings 2 Limited - MEIF II CP Holdings 1 Limited - MEIF II CP Holdings SARL - Macquarie European Infrastructure Fund II.

4.5 Blurring nationality

In Figure 7, we assess the relevance of peculiar cases of foreign subsidiaries, when the identification of nationality of the company is not trivial, because control paths can cross national borders. The problem is well known by international agencies promoting standards for foreign direct investment (e.g. UNCTAD, 2016). they may have important implications at the moment of drafting investment policies that want to discriminate based on investor nationality (UNCTAD, 2016).

Figure 7: Special cases: indirect foreign, transit investment, and round-tripping



As from our elaborations, we find that 24.51% of foreign subsidiaries in our data are actually only *indirect foreign*. That is, they belong to an *ultimate* parent company abroad, but indirect control is exerted through at least one domestic *middleman* subsidiary. We report the estimates of this case from our data in the first panel of Figure 7. On the second panel of the same figure, we report companies with *multiple passports*²³, i.e. subsidiaries whose control path involves more than one country before ending in the country of origin of the parent. They account for 19.16% of the total of foreign affiliates in our data. A third peculiar case is reported in panel c) of Figure 7. A *round-tripping* investment occur when the investor brings capital abroad in an incorporated company, then she starts from there an operation that brings the capital back in her country of origin as direct investment in another company. From our data, only 1.33% of foreign subsidiaries can actually be considered round-tripping investment operations from our data. Main factors explaining round-tripping may include the possibility to exploit offshore advantages in tax heavens, as well as a scant protection of property rights in the origin country (see for example Ledyeva et al., 2015, on the Russian case), which requires shielding capital abroad .

Indirect foreign subsidiaries and subsidiaries with multiple passports²⁴ are however relatively neglected in existing literature, although they seem rather relevant from our data.

²³Extracting randomly from our data, consider the case of Safilo Optical Sdn Bhd in Myanmar. It is finally controlled by the Safilo Group SpA in Italy, although we find *middlemen* first in Singapore and then in Hong Kong.

²⁴In a recent document, OECD (2015) suggests reporting separately the case of special purpose entities (SPEs), which channel investments (i.e. as *middlemen*) through several countries. Hence, SPEs are defined as entities that have little or no employment, physical presence, or operations in a country, while holding assets and liabilities or raising capital for the multinational enterprise. Actually, a cursory glance to our data suggests that only few *middlemen* subsidiaries may be alleged SPEs, while most have non-negligible employment and may be actively engaged in economic activities in

5 Empirical results

We organize econometric investigations as follows. First, we test the emergence of indirect *vis à vis* direct control on a subsidiary, as the first choice is the way parent companies start organizing pyramidal structures. In particular, we want to assess whether a parent’s choice is driven by institutional environment and geography.

Second, once pyramidal structures emerge, we test whether institutions and geography matter for running control paths between a specific couple of countries of origin and final destination. Finally, we test the choice of intermediate countries, as jurisdictions where *middlemen* subsidiaries are located.

5.1 Emergence of indirect control

We assume that a parent company can choose either to establish direct control over a subsidiary or to coordinate the latter through one or more *middlemen* subsidiaries, i.e. through a corporate control path with one or more predecessors in corporate control, following eq. 6. Eventually, if indirect control is chosen, a parent embeds that subsidiary in its pyramidal structure.

Accordingly, we test the following probit model:

$$Indirect_{h_0(c_o)i(c_i)} = \beta_0 + \beta_1 X_{c_o} + \beta_2 X_{c_i} + \beta_3 Z_{c_o c_i} + \beta_4 F_{h_0} + \beta_5 F_i + \varepsilon_{h_0(c_o)i(c_i)} \quad (12)$$

where the dependent variable is binary and equal to one if the parent company chooses indirect control on a subsidiary, and zero otherwise. The sets of regressors, X_{c_o} and X_{c_i} , include characteristics of the country of origin of the parent and of the subsidiary, respectively, with a focus on financial and contractual institutions. Among them, both metrics of financial development and contract enforcement are sourced from the World Bank’s Development Indicators and relative to year 2015. The first is the value of domestic credit provided to the private sector as a percentage of GDP. The second is the cost required to enforce a contract through the courts, as a percentage of the claim. Both measures have been used frequently in other works to assess the quality of financial institutions (see, for example, Acemoglu et al., 2009, and Rajan and Zingales, 1998b). Other country level controls have been sourced from World Bank, relative to year 2015. Specifically, we control for the entry cost required to start a business in the country of a subsidiary, as a percentage of income per capita. Further, we include tax rates on commercial profits and levels of GDP for both countries of parents and subsidiaries. Additionally, the set $Z_{c_o c_i}$ collects bilateral gravity controls sourced from Head et al. (2010) and Head and Mayer (2013), as updated to year 2015, with a specific focus on geography, economic and commercial agreements, and common institutions between parent’s and subsidiaries’ locations. Firm size categories²⁵ (small, medium, large, very large) are included in F_{h_0} and F_i . Standard errors are clustered by parent company. We report nested results in Table 4.

We find that it is more likely a parent chooses direct control in presence of good institutions in its origin country, adopting a more transparent governance for its subsidiaries, specifically in the case of industrial groups.²⁶ On the contrary, lower financial and contractual frictions in the location of subsidiaries more likely foster choices of indirect control, probably because parents are more able to coordinate from remote economic activities of subsidiaries. In this context, it seems that institutions act as a *push & pull* factor. Depending on which side of the control paths we look at: a) top-down, financial frictions encourage direct control paths, from the parent’s perspective, when we look at the origin countries; b) bottom-up, they encourage indirect control paths, when we look at hosting countries of subsidiaries.

Our interpretation of results in Table 4 find confirmation in analyses of following sections, when we study the *within* parent choice of preferred destinations of indirectly controlled subsidiaries, and then

their hosting countries.

²⁵See note 16 above for details on how categories are originally drawn in the Orbis database.

²⁶Based on the activity of the parent, we consider as industrial the groups of subsidiaries that are not led by financial companies: banks, insurance companies, private equity, mutual and pension funds.

Table 4: Emergence of indirect control

Dependent variable:	(1)	(2)	(3)	(4)
Indirectly controlled _{h0(c0)(c)}	All	All	Industrial group	Financial group
Financial development _{ct0}	.0006** (.0002)	-.0008*** (.0002)	-.0007*** (.0002)	-.0001 (.0006)
Financial development _{cti}	.0031*** (.0001)	.0026*** (.0001)	.0030*** (.0001)	.0005* (.0003)
Entry cost _{cti}	.0020** (.0007)	.0012 (.0007)	.0017** (.0009)	-.0031** (.0015)
Contract enforcement _{ct0}	-.0027** (.0012)	-.0020 (.0012)	-.0042*** (.0014)	.0008 (.0030)
Contract enforcement _{cti}	.0054*** (.0004)	.0057*** (.0004)	.0059*** (.0005)	.0046*** (.0010)
Profit tax _{ct0}	-.0026** (.0011)	-.0029** (.0011)	-.0044*** (.0012)	.0052* (.0024)
Profit tax _{cti}	.0018*** (.0004)	.0014*** (.0004)	.0014*** (.0004)	.0012 (.0008)
(log of) GDP _{ct0}	.0565*** (.0065)	.0642*** (.0064)	.0600*** (.0067)	.0437** (.0192)
(log of) GDP _{cti}	.0264*** (.0038)	.0381*** (.0040)	.0346*** (.0045)	.0650*** (.0097)
(log of) geographic distance _{cti}		-.0185* (.0098)	-.0193* (.0102)	.0412* (.0243)
Contiguous borders _{cti}		-.2443*** (.0190)	-.2700*** (.0213)	-.1533*** (.0346)
Common language _{cti}		.0411** (.0175)	.0610*** (.0200)	-.1533*** (.0346)
Colonial relationship _{cti}		.0823*** (.0225)	.0824*** (.0254)	.0967* (.0513)
Common currency _{cti}		.1818*** (.0190)	.1728*** (.0205)	.1054*** (.0437)
Common legal origin _{cti}		-.0985*** (.0142)	-.1054*** (.0154)	-.0679** (.0323)
WTO members _{cti}		-.0594 (.1032)	-.0216 (.0999)	-.2447 (.2230)
Regional trade agreement _{cti}		.0559*** (.0178)	.0252 (.0190)	.0439 (.0473)
Constant	-2.4947*** (.1565)	-2.7379*** (.1808)	-2.5817*** (.1846)	-2.9370*** (.5245)
Pseudo R squared	.0251	.0293	.0329	.0273
Log pseudolikelihood	-406,283.88	-404,562.31	-320,718.85	-74,199.48
Errors clustered by parent	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
N. of observations	604,785	604,785	478,432	126,353

Clustered standard errors by parent company in parenthesis ***

p<0.01, ** p<0.05, * p<0.1

its choice of intermediate jurisdictions where to locate *middlemen* subsidiaries. In Appendix Table B3, we try a robustness check limiting the analysis on domestic subsidiaries, for which the host country is the same of the parent company. We find that the effect of financial frictions on the choice of indirect control is statistically significant only in the case of domestic companies owned by MNEs, whereas no significance is detected in the case of domestic parent companies controlling exclusively domestic subsidiaries. The *push & pull* effect of financial frictions manifest itself when companies can choose among different institutional environments, at home and abroad, along corporate control paths.

As expected, lower tax rates in the country of the parent discourage the establishment of indirect control on any subsidiary, allowing for more transparency in governance. On the contrary, higher tax rates in the country of subsidiaries incentive indirect corporate control. Indeed, we can reasonably suspect that pyramidal structures, through indirect control, are established also to shift and hide profits when taxation is too high.

Among other controls, we find that geographic contiguity and longer geographic distances correlate with direct control. On the other hand, it is more likely that bigger countries sharing a common

language, a common currency and with a colonial common past establish indirect corporate control relationships.

In the last column of Table 4, we separate financial groups that are led by financial parents (banks, insurance, mutual funds, etc.). As expected, in this case, results are not robust and significant on institutions and taxation, possibly because their corporate governance is mainly based on the possibility to exploit financial gains, rather than on the necessity to coordinate productive activities.

5.2 Origin and destination of indirect control paths

From this section we restrict our analyses only to cases of indirect control in MNEs, to test which combinations of countries of origin and destination prevails, after looking at institutional environments, physical and political geography.

In the following equation 13, the dependent variable is an integer counting of how many times a parent company h_0 located in c_0 has made the choice to locate subsidiaries in a country c_i . It measures the within choice of a parent company to pick a destination, given its original location.

$$N_{h_0(c_0c_i)} = \beta_0 + \beta_1 X_{c_0} + \beta_2 X_{c_i} + \beta_3 Z_{c_0c_i} + \beta_4 F_{h_0} + \varepsilon_{h_0(c_0c_i)} \quad (13)$$

Based on the nature of the dependent variable, we implement a negative binomial regression model. On the right-hand side of the equation, we test the same determinants as in eq. 12, now excluding only the firm size of the subsidiary, because now the information is more aggregated. Standard errors are clustered by parent company. Nested results are reported in Table 5.

In this case, we find that what matters is exclusively the level of financial development of the country where the subsidiary is located, after a parent has chosen to embed the subsidiary in a pyramidal structure. In presence of lower financial frictions, more indirectly controlled subsidiaries will be located in a hosting country. This finding confirms that the parent establishes indirect control preferably where it is easier to coordinate management decisions from remote. However, in this case contract enforcement is to be not statistically significant as a driver for location of subsidiaries.

Finally, we observe that the choice of a negative binomial model over a simpler Poisson model was valid since the estimated dispersion parameter is significantly above one.

5.3 Indirect control and the role of intermediate countries

At this stage, we test which features, among institutions and geography, drive the parent's choice to choose an intermediate jurisdiction along indirect control paths. We estimate the following equation with a negative binomial regression model:

$$N_{h_0(c_0c_m c_i)} = \beta_0 + \beta_1 X_{c_m} + \beta_2 Z_{c_0c_m c_i} + \beta_3 F_{h_0} + \theta_{c_0} + \lambda_{c_i} + \varepsilon_{h_0(c_0c_m c_i)} \quad (14)$$

where the dependent variable is an integer indicating the number of subsidiaries indirectly controlled by a parent company h_0 originated in the country c_0 , which are organized on control paths crossing (one or more) intermediate countries c_m , before reaching destination country c_i . This time, we include a set X_{c_m} on the right-hand side that is made of country indicators for intermediate countries. Then, we add a set of gravity indicators, $Z_{c_0c_m c_i}$, for the multilateral characteristics of the countries involved in the entire control path.

Table 5: Origin and destination of indirect control paths, geography and institutions

Dependent variable:	(1)	(2)	(3)	(4)
Number of indirect control paths $_{Sh0(c0 ci)}$	All	All	Industrial group	Financial group
Financial development $_{c0}$.0001 (.0013)	-.0001 (.0014)	.0016 (.0012)	-.0025 (.0023)
Financial development $_{ci}$.0049** (.0012)	.0033*** (.0011)	.0049*** (.0010)	.0002 (.0016)
Entry cost $_{ci}$	-.0090 (.0064)	-.0088 (.0059)	-.0099* (.0058)	-.0115* (.0062)
Contract enforcement $_{c0}$	-.0019 (.0069)	-.0037 (.0074)	.0003 (.0054)	-.0093 (.0101)
Contract enforcement $_{ci}$.0013 (.0043)	.0015 (.0047)	.0057 (.0044)	-.0036 (.0049)
Profit tax $_{c0}$.0002 (.0008)	-.0025 (.0077)	.0010 (.0061)	-.0060 (.0117)
Profit tax $_{ci}$	-.0080* (.0041)	-.0100*** (.0033)	-.0054 (.0037)	-.0124*** (.0036)
(log of) GDP $_{c0}$	-.0969** (.0421)	-.0685* (.0385)	-.0784** (.0342)	-.0373 (.0521)
(log of) GDP $_{ci}$.2861*** (.0312)	.3204*** (.0303)	.3042*** (.0345)	.2782*** (.0329)
(log of) geographic distance $_{c0ci}$.0749 (.0490)	.0349 (.0663)	.1949** (.0771)
Contiguous borders $_{c0ci}$		-.1176 (.1644)	.0557 (.1988)	-.2954* (.1689)
Common language $_{c0ci}$.0650 (.1330)	.0173 (.1632)	.2240 (.1403)
Colonial relationship $_{c0ci}$.3650** (.1765)	.1804 (.1942)	.8756*** (.2247)
Common currency $_{c0ci}$.3615* (.2005)	.1713 (.1597)	.6251** (.2447)
Common legal origin $_{c0ci}$		-.0312 (.1370)	.1968 (.1360)	-.5262** (.2054)
WTO members $_{c0ci}$		-.8252* (.4689)	-.0109 (.3540)	-1.1419** (.5301)
Regional trade agreement $_{c0ci}$		-.1954 (.1340)	-.0109 (.1573)	.6129*** (.1804)
Constant	-2.173** (1.0445)	-3.4750*** (.9367)	-4.0661*** (1.1201)	-2.6452* (1.4690)
Pseudo R squared	.0288	.0312	.0384	.0281
Log pseudolikelihood	-1,373,088.7	1,372,756.5	-986,059.9	-380,433.11
Estimated dispersion parameter	1.7819***	1.7547***	1.7004***	1.7128***
Errors clustered by parent	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
N. of observations	329,586	329,586	240,096	89,400

Clustered standard errors by parent company in parenthesis ***

p<0.01, ** p<0.05, * p<0.1

All characteristics, X_{c_m} , of intermediate jurisdictions are averaged when there is more than one country along the control path. That is, we calculate their average financial development, their average entry cost, their average contract enforcement, and so on. As a novelty, X_{c_m} includes a variable that indicate whether on the path there is at least one country that can be considered an *offshore financial center*, as following classification by international assessment.²⁷

Among multilateral controls, $Z_{c_0 c_m c_i}$, geographic distance is summed up to take into account the entire route running from the parent country c_0 , going through the intermediate countr(ies) c_m , and finally reaching a destination c_i . The rest of multilateral controls are binary variables equal to

²⁷We define a country an *offshore financial center* if it classified as such by both the IMF (2014) and the OECD (2016).

Table 6: Intermediate jurisdictions along control paths

Dependent variable:	(1)	(2)	(3)	(4)
Number of indirect control paths (c_0 - c_m - c_i)	All	All	Industrial group	Financial group
Financial development $_{cm}$.0061*** (.0020)	.0049*** (.0014)	.0028*** (.0010)	.0060*** (.0019)
Offshore financial country $_{cm}$.2666* (.1383)	.2975** (.1126)	.3924*** (.1395)	.2554** (.1356)
Entry cost $_{cm}$	-.0134 (.0165)	-.0069 (.0164)	-.0314*** (.0091)	.0059 (.0228)
Contract enforcement $_{cm}$.0028 (.0080)	.0034 (.0069)	.0135** (.0058)	-.0102 (.0097)
Profit tax $_{cm}$	-.0030 (.0061)	-.0054 (.0044)	-.0012 (.0048)	-.0106* (.0064)
(log of) GDP $_{cm}$.2298*** (.0416)	.2538*** (.0339)	.2847*** (.0362)	.2137*** (.0472)
(log of) geographic distance $_{c_0-c_m-c_i}$.2943*** (.0460)	.1806*** (.0448)	.2137*** (.0472)
Contiguous borders $_{c_0-c_m-c_i}$.5415*** (.1735)	.4036* (.2199)	.7644*** (.1556***)
Common language $_{c_0-c_m-c_i}$.1599 (.2008)	.3024 (.2320)	-.2496 (.1603)
Colonial relationship $_{c_0-c_m-c_i}$.4185* (.2427)	.3365 (.2778)	.6831*** (.2184)
Common currency $_{c_0-c_m-c_i}$.4086** (.1645)	.3310** (.1315)	.5198** (.2272)
Common legal origin $_{c_0-c_m-c_i}$		-.1361 (.1014)	-.0168 (.1078)	-.2899** (.1359)
WTO members $_{c_0-c_m-c_i}$		-.2803* (.1595)	-.3230* (.1734)	-.3371* (.1838)
Regional trade agreement $_{c_m-c_0-c_i}$.4256*** (.1234)	.1420 (.1099)	.8589*** (.1871)
Constant		-6.5708*** (.8440)	-6.6205*** (.9475)	-6.5552*** (1.2572)
Pseudo R squared	.0202	.0292	.0348	.0302
Log pseudolikelihood	-1,667,385	-1,625,354.3	-1,167,943.1	-449,763.44
Estimated dispersion parameter	1.8469***	1.7477***	1.7011***	1.6815***
Errors clustered by parent	Yes	Yes	Yes	Yes
Origin-destination fixed effects	Yes	Yes	Yes	Yes
Firm level controls	Yes	Yes	Yes	Yes
N. of observations	397,645	390,883	285,406	126,353

Clustered standard errors by parent company in parenthesis ***
 $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

one if (all) intermediate countr(ies) on the control path verify a common condition, respectively, of contiguity, common language, shared colonial relationship, common currency, common legal origin, WTO membership, and they are all parts of a regional agreement.

The environment of countries of origin and destination is neutralized through the inclusion of two sets of fixed effects, θ_{c_0} and λ_{c_i} , respectively for the countries of the parent and of the subsidiary. Parent size is introduced under F_{h_0} . Standard errors are clustered by parent company. Nested results are reported in Table 6.

Interestingly, we find that also the level of financial development in intermediate jurisdictions is a significant driver of the parent choice in indirect control paths. This finding adds to previous results shown in Tables 4 and 5, where we commented that parent companies prefer to minimize coordination costs along control paths.

Nonetheless, *ceteris paribus*, we also find that a parent company prefers an *offshore financial center* to locate a *middleman* subsidiary, probably due to the possibility to limit disclosure of its financial information and/or to gain from favorable taxation. The (average) level of tax rates on profit in intermediate countries is not statistically significant. Finally, richer countries, more distant but

sharing a national border with the parent and the final subsidiaries, are preferably chosen for the location of *middlemen* subsidiaries.

6 Conclusions

As far as we know, our contribution is the first to study the ownership and corporate control of companies systematically, on a global scale, after adopting a basic network framework for webs of interlocking shareholding activities.

First, we propose a simple generalization of an ownership space, which we imagine as the common playfield for all investors in the share capital of firms. Then, we introduce a model to detect a concentration of voting rights iteratively, in the presence of pyramidal structures and cross-holding exchanges among companies. As a result, we can assign the role of parents and subsidiaries in ownership networks, while separating corporate control relationships from portfolio investment operations. We show that corporate boundaries can represent peculiar networks, i.e., *hierarchies of firms*, where parent companies on top can enforce management decisions in downstream subsidiaries, which organize themselves on different hierarchical layers of corporate control.

In general, we believe that our network framework is particularly useful to understand, on one hand, different shareholding architectures around the world and, on the other hand, further dimensions involved in the design of corporate boundaries, within and across countries.

In this context, we stress the role of *middlemen subsidiaries*, which are crucial in the organization of pyramidal structures. These are companies that are controlled by a parent on top but in turn control *subsidiaries of subsidiaries* on downstream lower layers of the hierarchy. As far as we know, existing economic theory and empirics neglect both the emergence of indirect corporate control, as well as the role of middlemen subsidiaries and their intermediate jurisdictions.

After we exploit a unique dataset of more than 53.5 million of firms, we provide some stylized facts on the heterogeneous distributions of corporate control networks, *aka* hierarchies of firms, *aka* corporate boundaries, both within and across countries. After that, we test how financial and contractual frictions can affect the emergence of indirect control relationships, as well as the geographic directions of these relationships. Fundamentally, we uncover a *push & pull* effect started by institutional environments in MNEs structures. Less financial and contractual frictions in the country of a parent drive to more transparent forms of corporate governance, whereas less financial frictions in the countries of subsidiaries, including *middlemen*, permit a stretching of pyramidal structures and the establishment of indirect control, possibly because a parent can better afford to coordinate management decisions from remote.

Ceteris paribus, we also detect a role for offshore financial centers, as these may be preferred intermediate jurisdictions for their lack of financial disclosure and for their lower level of taxation.

We believe that our work paves the way to a variety of follow-ups. Once the network nature of shareholding and control is known, we can study for example how the intra-firm trade of goods and services develops on this network, probably at prices commanded by the headquarters, which are different from market prices, initiating profit-shifting operations. Else, we can study if and how corporate networks can establish internal capital markets, switching resources across countries and industries on unsynchronized business cycles, as an alternative to costly resources on external financial markets. More interesting, it seems to us, is the possibility to understand if and how the concentration of corporate control observed here also implies an increasing concentration of market power, raising an issue of lack of competition that can be detrimental to overall welfare.

References

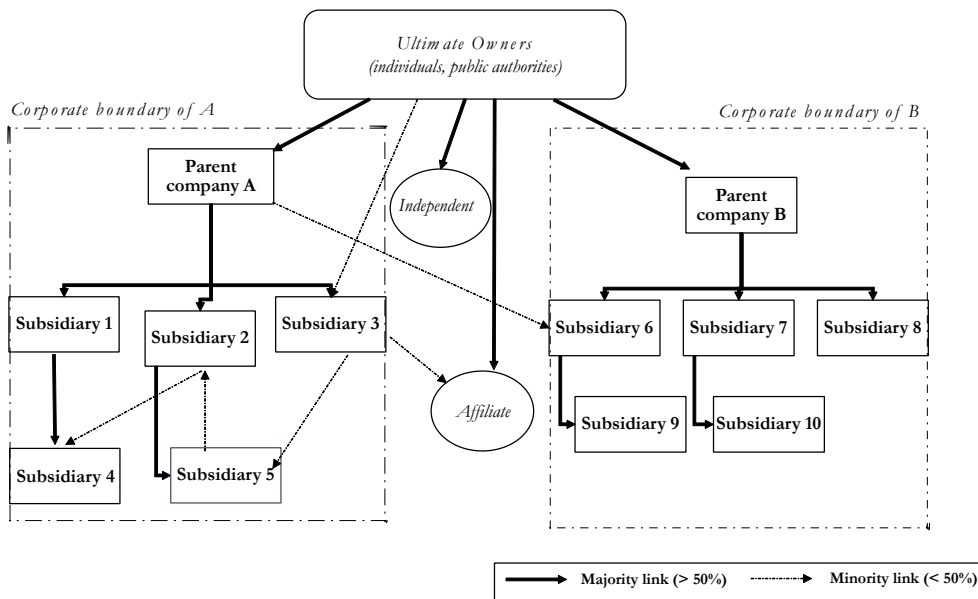
- Acemoglu D., Johnson S. and Mitton T. (2009). 'Determinants of Vertical Integration: Financial Development and Contracting Costs,' *Journal of Finance* vol. 64(3), pp. 1251-1290.
- Almeida H. and Wolfenzon D. (2006). 'A Theory of Pyramidal Ownership and Family Business Groups,' *Journal of Finance*, vol. 61 pp. 2637-2681.
- Altomonte C. and Rungi A. (2013). 'Business Groups as Hierarchies of Firms: Determinants of Vertical Integration and Performance,' *European Central Bank Working Paper Series N.1554*.
- Banzhaf J.F. (1965). 'Weighted voting does not work: a mathematical analysis,' *Rutgers Law Review*, vol. 19 pp. 317-343.
- Berglof E. and Perotti E. (1994). 'The Governance Structure of the Japanese Financial Keiretsu,' *Journal of Financial Economics*, vol. 36 pp. 259-284.
- Berle A. A. and Means G. C., (1967). 'The Modern Corporation and Private Property', second revised edition. Brace and World, New York.
- Bohren, O. and D. Michalsen. 1994. "Corporate Cross-Ownership and Market Aggregates: Oslo Stock Exchange 1980-1990." *Journal of Banking and Finance*, vol. 18, no. 4: 687-704.
- Chapelle A. and Szafarz A. (2007). 'Control consolidation with a threshold: an algorithm,' *IMA Journal of Management Mathematics* vol. 18 pp. 235-243.
- Claessens S., Djankov S. and Lang L. H. P. (2000). 'The separation of ownership and control in east Asia corporations,' *Journal of Financial Economics*, vol. 58 pp. 81-112.
- Crama Y. and Leruth L. (2013). 'Power Indices and the Measurement of Control in Corporate Structures,' *International Game Theory Review*, vol. 15 (3).
- Dubey P. and Shapley L. (1979) 'Mathematical Properties of the Banzhaf Power Index', *Mathematics of Operations Research*, Vol. 4 (2) pp. 99-131.
- Faccio M., Lang L. H. P. (2002). 'The Ultimate Ownership of Western European Corporations,' *Journal of Financial Economics*, vol. 65 pp. 365-395.
- Fama E. F. and Jensen M. C. (1983). 'Separation of Ownership and Control,' *Journal of Law and Economics*, vol. 26(2) pp. 301-325.
- Head K., Mayer T. and Ries J. (2010). 'The Erosion of Colonial Trade Linkages after Independence', *Journal of International Economics* vol. 81(1) pp.1-14.
- Head K. and Mayer T.(2013). 'Gravity Equations: Toolkit, Cookbook, Workhorse'. In *Handbook of International Economics*, Vol. 4, eds. Gopinath, Helpman, and Rogoff, Elsevier Publishing.
- IMF (2014) *Offshore Financial Centers (OFCs): IMF Staff Assessments (OFCA)*. International Monetary Fund. IMF, New York.
- Jacomy M, Venturini T, Heymann S and Bastian M (2014). 'ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software', *PLoS ONE* 9(6).
- La Porta R., Lopez-De-Silares F., and Shleifer A. (1999). 'Corporate Ownership Around the World,' *The Journal of Finance*, vol. 54(2) pp. 471-517.
- Ledyeva S., Karhunen P., Kosonen R., and Whalley J. (2015). 'Offshore Foreign Direct Investment, Capital Round-Tripping, and Corruption: Empirical Analysis of Russian Regions,' *Economic Geography* vol. 91 (3), pp. 305-341.
- Levy M. (2011). 'The Banzhaf Index in complete and incomplete shareholding structures. A new algorithm,' *European Journal of Operational Research*, vol. 215 (2) pp. 411-421.
- Levy M. and Szafarz A (2017). 'Cross-Ownership: A Device for Management Entrenchment?', *Review of Finance*, vol. 21 (4) pp 1675-1699.
- Mathews R.D. (2007). 'Optimal Equity Stakes and Corporate Control,' *Review of Financial Studies*, vol. 20 pp. 1059-1086.
- Means G. C. (1931), 'The Separation of Ownership and Control in American Industry', *The Quarterly Journal of Economics*, vol. 46 (1), pp. 68-100.
- Morck R., Wolfenzon D. and Yeung B. (2005), 'Corporate Governance, Economic Entrenchment, and Growth', *Journal of Economic Literature*, vol. 43 pp. 655-720.
- Nenova T. (2003). 'The Value of Corporate Voting Rights and Control: A Cross-country Analysis', *Journal of Financial Economics*, vol. 68 (3) pp. 325-351.
- OECD (2015a). *Measuring Globalisation: OECD Handbook on Economic Globalisation Indicators*. OECD Publishing.

- OECD (2015b). 'FDI Statistics by Ultimate Investing Country'. In *Implementing the Latest International Standards for compiling Foreign Direct Investment Statistics*. OECD Publishing.
- OECD (2016). *Report to G20 Leaders*. OECD Publishing.
- OECD (2017). 'Cross-border M&A on the Rise', background note at the Global Forum on International Investment. OECD, Paris.
- Rajan R. and Zingales L. (1998a). 'Power in a theory of the firm,' *Quarterly Journal of Economics*, vol. 113 pp. 387–432
- Rajan R. and Zingales L. (1998b). 'Financial dependence and growth', *American Economic Review* vol. 88, pp. 559–586.
- Shapley L. S. (1953). 'A Value for n-person Games'. In Kuhn, H. W.; Tucker, A. W. *Contributions to the Theory of Games*. *Annals of Mathematical Studies*, pp. 307–317. Princeton University Press.
- UNCTAD (2009). *The UNCTAD Manual on Statistics of FDI and the operations of TNCs*. United Nations, Geneva.
- UNCTAD (2016). *World Investment Report. Investor Nationality: Policy Challenges*. United Nations, Geneva.
- UNCTAD (2017). *World Investment Report. Investment and the Digital Economy*. United Nations, Geneva.
- Vermeulen E. (2013). 'Beneficial Ownership and Control: A Comparative Study - Disclosure, Information and Enforcement', *OECD Corporate Governance Working Papers*, No. 7, OECD Publishing.

APPENDIX A: Definitions

In Figure A1 we sketch a fictional ownership network, where we can spot two different corporate boundaries, *aka* corporate networks. Assuming we were able to separate corporate boundary A from corporate boundary B following Section 3, we can summarize here the definitions of Ultimate Owners, parent company, subsidiary company, affiliate, and independent company that we use throughout our analyses.

Figure 8: A stylized ownership network with two corporate boundaries.



Ultimate owners (UOs) are the shareholders (individuals, families, public authorities) that are on top of any ownership path, because they cannot be owned by any other shareholder in nature or by law. We can also say that they are the *sources* of the ownership space introduced from Section 2, because they do not have any *predecessor* in ownership. In fact, they are the starters of any ownership network as from initial decisions to invest in companies' capital shares. As depicted in Figure A1, we consider UOs as on top but outside of corporate boundaries. See also Figure 3.

Our loose notion of UOs provided here does differ from a more specific legal definition of Ultimate Beneficial Owner (UBOs), which identifies subjects that are for their actions ultimately liable in front of the law. For example, the EU's Fourth Anti-Money Laundering Directive (MLD4) identifies an UBO as a subject that holds more than 25% of voting rights in a legal entity. On the other hand, the Financial Action Task Force on Money Laundering (FATF) is an intergovernmental body participated by 37 member countries, which defines an UBO more flexibly as the natural person who ultimately own or control a legal entity and/or the natural person on whose behalf a business is being conducted.

We argue that keeping our definition of UOs more inclusive allows us to catch better the variety of wedges that can be inserted between ownership and control of a company, given also the variety of corporate structures we may encounter (on this issue, see also Vermeulen et al., 2013).

A **parent company** is a company that coordinates the activities of one or more subsidiary company. It is not controlled by any other corporate shareholder, although one or more corporate shareholders can sit at its assembly with no dominant stake.

We can consider exceptions the cases of parent companies whose dominant stakes are in the hands of financial intermediaries, like mutual funds or pension funds, which can operate as shareholders but on behalf of their individual investors. In fact, the latter are cases where a distinction between

ownership and corporate control can be murky, as it is based on the more active or passive role played by institutional investors across different national jurisdictions and in the different moments of life of a company. See also Section 3.2.2 for further discussion of peculiar dominant stakes in light of a control probability framework.

A **subsidiary company** is a company that is directly or indirectly controlled by a parent company through ownership linkages. In the case of direct corporate control, the absolute majority of capital shares of a subsidiary is held by a parent company. In the case of indirect corporate control, one or more subsidiaries already controlled by a parent company in turn control also a '*subsidiary of subsidiary*'. Indirect control implies the arrangement of subsidiaries on hierarchical layers starting from the parent company, as in Figure A1.

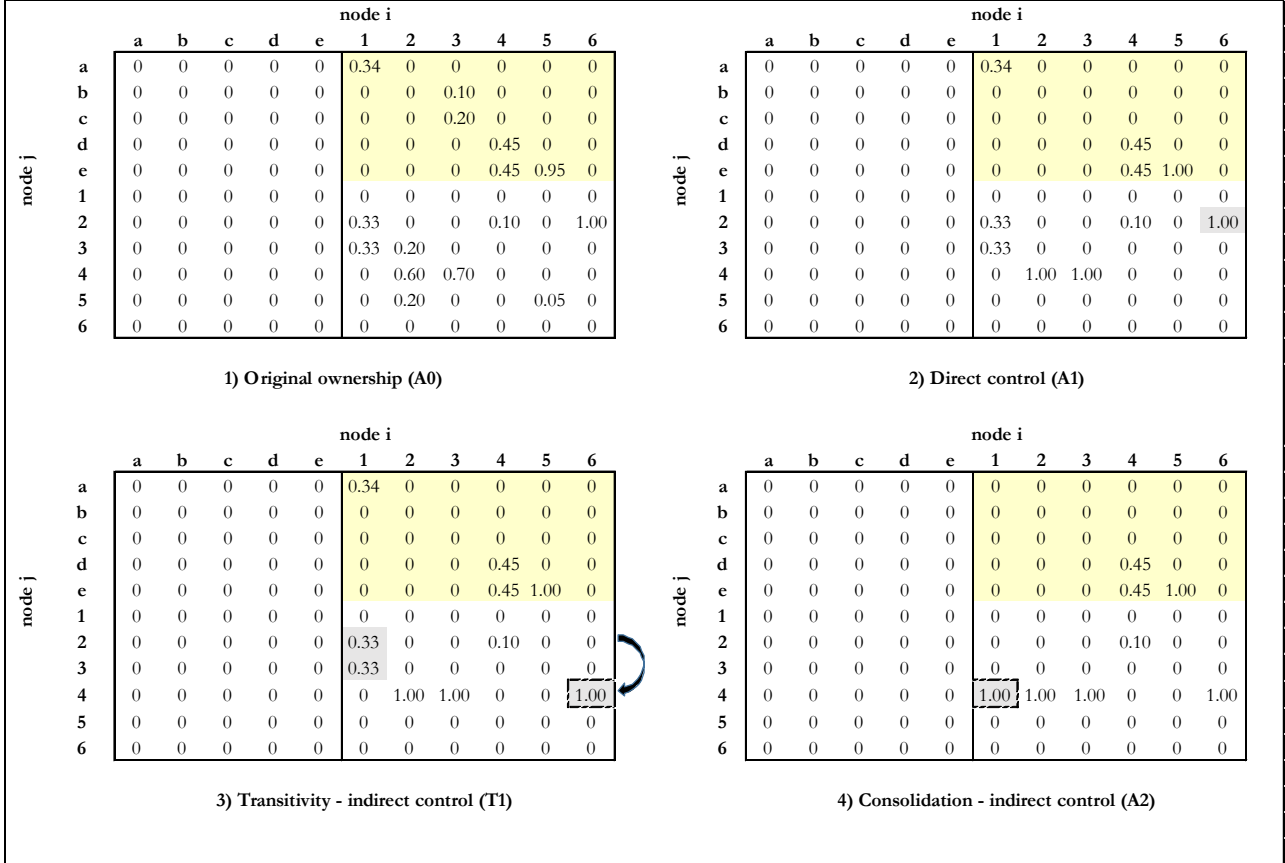
An **affiliated company** is a company whose parent only possesses a minority stake in the ownership of the company, while an **independent company** has no parent company at all.

Often it is a matter of degree of control, when ownership is fragmented, whether the balance tilts to a parent company or to one or more Ultimate Owners. In Figure A1 we located an affiliated company outside corporate boundary A, although we could consider it as included within a corporate boundary when the parent (or one of its subsidiaries) is influential enough on its management decisions, after a dominant stake. See also UNCTAD (2009) and OECD (2005) on international standards proposed for national statistics about foreign firms.

We believe that our probabilistic approach, after a Banzhaf Index is introduced in Section 3.2.2, allows to catch every shade from outright corporate control by one parent to the case of independent companies, when no unique controlling shareholder is spotted. Indeed, it includes cases when more than one parent company can claim equal control power on an affiliated company, because for example the affiliated company is on purpose established as a *joint venture* with equal voting rights shared by investors.

APPENDIX B: TABLES AND GRAPHS

Figure B1: From the ownership matrix to the control matrix. an application



Starting from the original ownership adjacency matrix (A0) of the illustrative network of Figure 3, any j th node may have a stake in an i th node. The coloured area is the sub-matrix of ultimate ownership, which does not change after application of the corporate control transformation. Following matrices A1, T1 and A2 represent applications of eqs. 7, 8 and 9, respectively. Grey-colored cells show substitutions. No iteration is needed in this case: matrix A2 represents corporate control in the given ownership network.

Figure B2: Banzhaf (1965) Index, capital share and number of shareholders for minority-owned companies

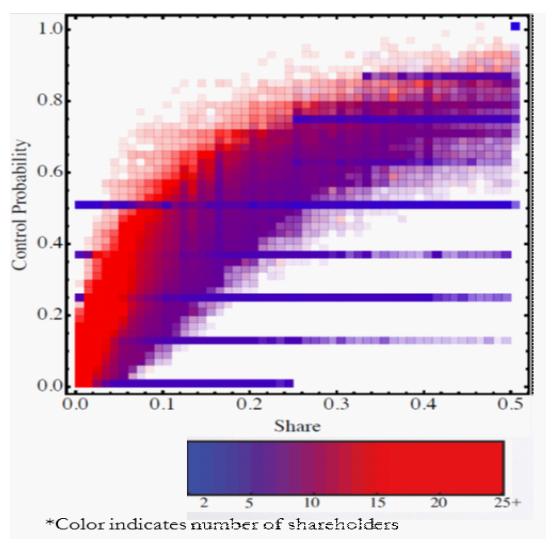


Table B1: Companies participating to ownership networks: geographic coverage and firm size

Economy	Small	Medium	Large	Total
European Union	2,920,807	713,176	323,924	3,957,907
of which:				
<i>Germany</i>	536,578	138,180	50,170	724,928
<i>France</i>	101,064	76,576	38,558	216,198
<i>United Kingdom</i>	437,778	77,485	68,046	583,309
<i>Italy</i>	203,179	96,872	32,245	332,296
<i>Spain</i>	133,862	48,305	24,021	206,188
United States	3,620,543	45,067	62,494	3,728,104
Russia	437,180	105,822	34,367	577,369
Other Europe	228,161	64,132	20,802	313,095
Asia	553,644	116,435	131,578	801,657
of which:				
<i>Japan</i>	118,195	27,931	30,774	176,900
<i>China</i>	99,942	10,874	34,595	145,411
<i>India</i>	17,032	4,738	8,990	30,760
Africa	39,750	2,151	4,475	46,376
Central and South America	121,580	20,159	17,752	159,491
of which:				
<i>Brazil</i>	8,631	9,534	5,365	23,530
<i>Argentina</i>	5,358	1,174	1,432	7,964
<i>Mexico</i>	12,994	1,754	2,627	17,375
<i>the Caribbean countries</i>	41,662	141	1,676	43,479
Australia	381,878	48,311	10,211	440,400
Rest of the world	606,339	21,300	17,473	645,112
TOTAL	8,909,882	1,136,553	623,076	10,294,391

Table B2: Top 20 countries where firms establish reciprocal cross-holdings

Top 20 countries	N. firms	Avg share	st. dev.	Min	Max
Germany	4956	92.69	23.17	0.03	100
United States	1556	50.09	8.59	0.10	100
Japan	1493	14.09	13.24	0.01	100
Russia	1427	36.97	36.89	0.01	100
Spain	1123	49.77	37.26	0.01	100
United Kingdom	1016	72.29	39.19	0.01	100
Australia	960	58.94	38.85	0.01	100
Belgium	787	43.06	41.16	0.01	100
Italy	648	29.73	33.44	0.01	100
Israel	630	43.91	43.49	0.01	100
Portugal	587	33.99	34.19	0.02	100
Ireland	293	49.36	39.43	0.01	100
France	207	28.59	35.79	0.01	100
Turkey	162	26.01	31.09	0.02	100
Norway	146	42.11	39.60	0.15	100
Bulgaria	128	50.51	40.75	0.02	100
Romania	116	34.12	37.00	0.02	100
Taiwan	109	14.92	25.54	0.03	100
Netherlands	107	74.02	33.15	0.01	100
Poland	106	58.13	36.04	0.01	100
<i>Other countries</i>	<i>1437</i>	<i>39.45</i>	<i>28.70</i>	<i>0.01</i>	<i>100</i>

Table B3: Indirect vs direct corporate control in domestic subsidiaries

Dependent variable:	All domestic subs	Domestic subs of MNEs	Domestic subs of domestic groups
Indirectly controlled _{h0(c0)i(c0)}			
Financial development _{c0}	-.0006* (.0002)	-.0020*** (.0004)	.0004 (.0003)
Entry cost _{c0}	-.0120*** (.0021)	-.0327*** (.0033)	-.0014** (.0005)
Contract enforcement _{c0}	.0022* (.0009)	.0134*** (.0015)	-.0005 (.0009)
Profit tax _{c0}	.0037*** (.0009)	.0005 (.0002)	.0052*** (.0009)
(log of) GDP _{c0}	-.1505*** (.0058)	-.0114 (.0121)	-.1785*** (.0057)
Constant	2.2776*** (.1307)	-.6311** (.2761)	2.870*** (.1422)
Pseudo R squared	0.3746	0.1013	0.2981
Log pseudolikelihood	-964,026.55	-290,510.97	-649,983.6
Errors clustered by parent	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes
N. of observations	3,665, 804	474,897	3,190, 804

Clustered standard errors by parent company in parenthesis. *** p<0.01, ** p<0.05, * p<0.1