4-19-2013

Business Groups as Hierarchies of Firms: Determinants of Vertical Integration and Performance

Carlo Altomonte
Bocconi University and FEEM

Armando Rungi
University of Warsaw and FEEM, armando.rungi@unibocconi.it

Business Groups as Hierarchies of Firms: Determinants of Vertical Integration and Performance

Carlo Altomonte
(Bocconi University)

Armando Rungi
(University of Warsaw & FEEM)

March 18, 2013

Abstract

We explore the nature of Business Groups, that is network-like forms of hierarchical organization between legally autonomous firms spanning both within and across national borders. Exploiting a unique dataset of 270,474 headquarters controlling more than 1,500,000 (domestic and foreign) affiliates in all countries worldwide, we find that business groups account for a significant part of value-added generation in both developed and developing countries, with a prevalence in the latter. In order to characterize their boundaries, we distinguish between an affiliate vs. a group-level index of vertical integration, as well as an entropy-like metric able to summarize the hierarchical complexity of a group and its trade-off between exploitation of knowledge as an input across the hierarchy and the associated communication costs. We relate these metrics to host country institutional characteristics, as well as to the performance of affiliates across usiness groups. Conditional on institutional quality, a negative correlation exists between vertical integration and organizational complexity in defining the boundaries of business groups. We also find a robust (albeit non-linear) positive relationship between a group's organizational complexity and productivity which dominates the already known correlation between vertical integration and productivity. Results are in line with the theoretical framework of knowledge-based hierarchies developed by the literature, in which intangible assets are a complementary input in the production processes.

JEL classification: F23; L22; L23; L25; D24, G34

Keywords: production chains, hierarchies, business groups, financial development, property rights, vertical integration, corporate ownership, organization of production, productivity

1 Introduction

"The economics literature has not had much to say about non-standard organizational forms [...] now much discussed in the business and organizational literatures, including joint ventures, strategic alliances, networks, business groups, clans, and virtual organizations". [Baker et al., 2001b]

The emergence of Business Groups (BG) is traditionally considered a phenomenon typical of countries at an early stage of development: in order to circumvent market imperfections, firms with a formally autonomous legal status are put under a common control exerted by a parent entity in a network-like hierarchical organization of economic activities[1]. These organizational forms thus provide at the same time incentives to self-inforce promises of cooperation among units of production, given the control exerted by a common parent, without giving up the advantage (if and when necessary) of organizing activities within a market-like environment, since each affiliate maintains formal property rights on its production assets.

Under this general definition, also multinational enterprises (MNEs) can be considered as Business Groups, since one of their distinctive features is to organize legally autonomous affiliates spanning across different countries under the common management of a unique headquarter. The latter is confirmed by a cursory glance at the data: the top 100 largest multinational enterprises by international assets listed by UNCTAD (2011) are all organized as cross-border BGs, with an average of 330 affiliates in 64 countries each, and up to 10 different hierarchical levels of control.

In order to characterize the phenomenon of Business Groups (also encompassing MNEs under this general definition) and derive the ensuing implications in terms of determinants of a firm's boundaries, we map 270,374 headquarters controlling 1,519,588 affiliates worldwide in 2010, across all industries.[2] As we have individual balance sheet data for (most of) these firms, we are able to recover a total (unconsolidated) value added accruable to Business Groups of some 27.9 US$ trillion. Moreover, the largest BGs (headquarters with more than 100 affiliates worldwide) constitute less than 1% of groups in our sample, but account for 72% of the total value-added of BGs measured in the data. In a nutshell, in our data some 2,000 Business Groups worldwide account for around 20 US$ trillion of (unconsolidated) value added in year 2010.

In terms of trade flows, a reading of the US BEA (2012) data along the dimension of Business Groups reveals that at least 75% of US trade can be linked to firms organized as multinational BGs.[3] A similar exercise for France, where transaction- and firm-level data have been matched to the ownership

[1] Across geography and time, the different notions of *chaebol* in South Korea, *keiretsu* in Japan, *konzerne* in Germany all make reference to the idea of clusters of firms under common control. Khanna and Yafeh (2007) provide a survey of Business Groups' presence in emerging countries. Jones and Colpan (2010) or Fruin (2008) explore their importance in the early history of industrialized nations.

[2] Our source of data is ORBIS, a global dataset containing detailed balance sheet information for some 100 million companies worldwide. In addition, the database contains information on over 30 million shareholder/subsidiary links, when relevant. The database has been significantly expanded since 2009, with a better coverage of countries traditionally not well mapped such as Japan and the United States. More detailed information on the dataset, as well as its validation across countries, is discussed in Section 2.

[3] The US BEA (2012) reports that in 2009 foreign affiliates located in the United States accounted for 20.8 percent of the country's exports and 31.1 percent of imports of goods. At the same time, U.S. exports of goods associated to US multinationals were 54.7 percent of total exports of goods, while the similar figure for imports was 45.1 percent. As a result, 75.5 percent of total U.S. exports and 76.2 percent of total U.S. imports of goods in 2009 can be considered as Business-Group related.

structure of companies, reveals that some 65% of total French imports or exports can be reconducted to firms (domestic or foreign-owned) that are part of a Business Group structure (Altomonte et al., 2012).

Notwithstanding the economic relevance of BGs, the theory of the firm has been relatively silent on these organizational forms, with most authors implicitly assuming that these structures could be epitomized by a simple two-dimensional decision problem at a firm-level: whether to source intermediate inputs from within the firm or not, i.e. the vertical integration decision; and whether to locate an economic activity in the country of origin or abroad, i.e. the offshoring decision.

The vertical integration decision has been explored by a vast literature modelling incomplete contracts and firm boundaries, based on the seminal works of Williamson (1971, 1975, 1985), Grossman and Hart (1986) and Hart and Moore (1990). The offshoring decision, instead, has been theoretically studied among others by Grossman and Helpman (2002, 2003, 2004, 2005), Antràs (2003), Antràs and Helpman (2004, 2009).[4]

A common finding of this literature is that firm boundaries depend on institutional frictions. In particular, Acemoglu, Johnson and Mitton (2009) are the first to empirically investigate the combined impact of financial and contracting institutions on vertical integration decisions, finding vertical integration to be positively correlated with the interaction term between contracting institutions and financial frictions. From a slightly different perspective, Alfaro et al. (2011) find that similar levels of protectionism, hence trade institutions, imply also similar levels of vertical integration. Alfaro and Charlton (2009) investigate vertical FDI activities and find that these are not explained by host countries' comparative advantages, as affiliates tend to be rather proximate to parents both in vertical integration and skill content. Nunn (2007) or Nunn and Trefler (2008) provide instead an empirical support for the main tenets of the literature on the offshoring decision, relating the contracting environment of a supplier's inputs to the share of US imports that are intra-firm.

All these papers neglect however the possibility that the decisions on vertical integration or offshoring can be undertaken considering the organizational form of a Business Group, an option that generates a number of additional trade-offs in the definition of firms' boundaries.[5]

Consider for example the case of two ex-ante similar Business Groups present in our dataset: General Motors and Mitsubishi. Both groups have a century-old tradition in the production of motor vehicles in their own country of origin (the US and Japan). Moreover, in 2010 our data report that these two groups have a similar size, as they control 659 and 652 affiliates in 54 and 32 countries, respectively.[6] Still, when looking at industrial activities beyond motor vehicles, Mitsubishi is involved

[4]See among others the surveys by Holmstrom and Tirole (1989), Whinston (2001), Joskow (2005), Helpman (2006), Antràs and Rossi-Hansberg (2009), Aghion and Holden (2011).

[5]The only attempt we have found to explicitely model a theory of business networks is in Kali (1999; 2003). However, also in his approach Business Groups are the result of either a limited contract enforcement or imperfect capital markets, with their nature thus essentially reconducted to the 'dual' nature of firm boundaries, without mentioning the implications of firms'hierarchies. On the other hand, the literature on hierarchies has traditionally been confined to organizational issues stemming from managerial incentives *within*, rather than across, individual firms' boundaries (Bloom, Sadun and Van Reenen, 2012; Caliendo and Rossi-Hansberg, 2012; Marin and Verdier, 2009; Grossman and Helpman 2004). Alternatively, the issue has been considered as yet another aspect of firms' size in the finance literature (Acharya, Myers and Rajan, 2011; Rajan and Zingales, 2001a, 2001b; Kumar, Rajan and Zingales, 1999).

[6]Alfaro and Charlton (2009) also recall the GM case and enlist 2,248 'entities' belonging to the GM network in 1999, making however no difference between affiliates/subsidiaries and branches/divisions as we do (see infra). However some major events have occurred to GM since 1999. In 2005 the group conclusively sold its participations in electronics

in some ten lines of business (e.g. electronic products, aircraft, shipbuilding, petroleum products, chemical products, primary metals, food & beverages, bank and insurance, real estate), while GM beyond motor vehicles provides only financial services for its customers. Accordingly, the affiliates of Mitsubishi are able to provide a wider range of intermediate inputs to the group, with firms typically operating in 3 or 4 main different industries, whereas the affiliates of General Motors seem relatively more focused on one or two main intermediate activities. As a result, the degree of vertical integration is higher for Mitsubishi than GM. Crucially, however, Mitsubishi is significantly less complex in terms of organization, with a much flatter hierarchical structure (with no more than 3 levels of hierarchy within the group), while GM is characterized by a deeper (up to 8 levels) and more complex hierarchy of cross-participations in its affiliates. Moreover, we also find that the labor productivity of affiliates belonging to the hierarchically more complex GM group is on average significantly larger than the one of Mitsubishi's affiliates.

The latter evidence, showing that vertical integration choices are not independent from decisions on the organization of the hierarchy of firms across groups, is systematic and statistically significant across our sample once we control for institutional characteristics of the host countries. Also, the finding that higher levels of complexity in hierarchies, rather than vertical integration levels, are positively associated with the average productivity of affiliates operating within a given group (controlling for the location and the main activity of affiliates and headquarters) is sistematic in our data.

In exploring these issues, the paper thus provides a number of contributions to the literature. First, thanks to the richness of our dataset encompassing information on roughly 1,800,000 firms between headquarters and affiliates worldwide, we aim at providing a preliminary comprehensive picture of the phenomenon of Business Groups across the globe. Two thirds of our BGs are originated in OECD economies, whose headquarters own about 76% of affiliates worldwide. The ratio of foreign to domestic affiliates is smaller for groups originating from developing countries (around .3), since these countries have a relatively larger proportion of firms organized as domestic business groups (Khanna and Yafeh, 2007), while the ratio is highest for the US (.85), where Business Group structures tend to operate abroad rather than domestically.[7] We also find individual affiliates of BGs to be systematically larger, more productive and more capital intensive vs. firms not operating within these structures, in line with the findings of Atalay et al. (2012).

As a second contribution, we improve on the exercise of Acemoglu, Johnson and Mitton, (henceforth AJM, 2009) who have studied the determinants of vertical integration in a large cross-country dataset. With respect to their analysis, we refine the notion of vertical integration by nesting in the latter measure an I/O matrix which is specific to the group structure, thus being able to distinguish between group- and affiliate-level vertical integration. We find this distinction to be relevant in our data, as it allows for a better identification of the relationship between institutional characteristics and vertical integration measures. The intuition here is that estimating vertical integration in a sample of firms in which each BG's affiliate is considered as an independent firm, as the literature has done

production (Hughes Electronics, Electro-Motive) and in 2006 left to Toyota the control of Subaru, Suzuki and Isuzu. As a consequence of the industrial restructuring undertaken in 2009, GM has given up production of some brands (e.g. Pontiac, Oldsmobile) and the European division has almost completely dissolved, leaving only Opel in Germany in charge of the remaining activities.

[7]This finding is generally consistent with the idea that the boundaries of the firm should be larger in the presence of a poor institutional environment and thus higher transaction costs.

insofar, would miss the structural correlation in vertical integration linking affiliates of the same group, thus generating potentially biased results. This feature of firms' boundaries shaped as Business Group, with the ensuing implications, has been neglected insofar in the analyses on vertical integration.

A third contribution hinges on the literature on organization and hierarchies. Borrowing from graph theory, we develop a measure of complexity applicable to any hierarchical organization (including Business Groups), which is consistent with theoretical models of knowledge-based hierarchies (as for example in Caliendo and Rossi-Hansberg, 2012), where a trade-off can arise between the exploitation of knowledge as an intangible input and its communication along the hierarchy. The measure is retrieved as a variation of the node entropy of a hierarchical graph, and is continuous and additive in the number of levels. In our sample the measure is also Pareto-distributed across groups, in line with the previously mentioned concentration of economic activity in the largest (and organizationally more complex) groups. Throughout the paper we show how hierarchical complexity interacts with the (refined) measures on vertical integration, so as to derive a more precise characterization of the boundaries of Business Groups.

We use these novel data and refined metrics to generate a number of results. Consistently with the property rights theory of the firm, we find that better institutions lead to less vertical integration, both at the group and at the affiliate level. Moreover, BGs that have a high internal degree of vertical integration (between headquarter and affiliates) also tend to have relatively unspecialized (more integrated) affiliates. Interestingly, the affiliate and its group are at the margin less similar in terms of vertical integration in 'good' institutional environments, as a higher contract enforcement and/or a better financial development allow the single affiliate to specialize more, exchanging fewer inputs with coaffiliates and the parent. Moreover, conditional on the quality of institutions, a negative correlation arises between vertical integration and organizational complexity: for a given level of financial development, more specialized (less integrated) affiliates end up within more complex organizational structures. Contractual enforcement yields a similar trade-off, but less robust.

We also find that the positive relationship between vertical integration and affiliates' productivity emerging in our data is not robust to the inclusion of a group's organizational complexity, thus providing yet another piece of evidence on the importance of considering jointly vertical integration and organizational complexity decisions in assessing BGs. Our result of a positive correlation between organizational complexity and productivity within BGs' affiliates (controlling for vertical integration) complements the finding of Atalay, Hortacsu and Syverson (2012) in the case of US firms, according to which much of the correlation between a firm's performance and its vertical structure fades away when controlling for a generic measure of firm size. The result is also consistent with the economic rationale provided by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007), according to which best intangible assets (such as best managers, best managerial practices) can be shared in presence of a larger number of units of production (in our case more complex hierarchies) and hence their cost can be smoothed on a larger scale.

The relevance of intangible assets is also confirmed by the fact that the relationship between organizational complexity and productivity is non-linear: above a certain threshold of complexity (around 550 affiliates and/or 5 levels of control) the relationship becomes negative. This result is in line with the microfoundation provided by Caliendo and Rossi-Hansberg (2012), in which a minimum efficient scale exists in the acquisition and communication of knowledge throughout the hierarchy, associated however to the emergence of endogenous communication costs of additional management layers, which should increase with complexity. Such an evidence of marginally decreasing returns from increasing complexity is relevant, as it puts a natural limit to the growth in complexity of business groups: indeed, only 1% of groups in our sample exceed this average 'optimal' organizational threshold.

The same theoretical framework of Caliendo and Rossi-Hansberg (2012) can also explain another result we have obtained when distinguishing between hierarchical complexity (which takes into account the overall density of affiliates at each level of the control hierarchy) and the simple hierarchical distance, i.e. the length of the command chain linking each affiliate to the parent company. In general, we find that the further the firm is from the decision making center, the lower its level of productivity appears to be; however, the latter result only holds when we control for our main measure of organizational complexity. When considering only hierarchical distance in the model (itself a raw proxy of organizational complexity), affiliates located at further levels of control would actually display higher levels of productivity. Our data are thus consistent with the idea that further layers of management allow for the exploitation of economies of scale for knowledge inputs, and hence affiliates to bigger (more complex) networks are relatively more productive. At the same time, subsidiaries located at further hierarchical distances from the headquarter discount a higher cost of communication and show (at the margin) a negative productivity premium.

One final caveat is worth mentioning: throughout the analysis we have explored the boundaries of Business Groups with respect to the 'make or buy' (vertical integration) decision, which we have then interacted with the extent of hierarchical complexity of the same group. We have instead considered as given the decision on whether to locate production at home or abroad, thus encompassing multinational enterprises as a subset of Business Groups. Hence, while in the analysis we always control for the potentially different behavior of foreign vs. domestic affiliates, in this paper we do not explicitly model the drivers of a foreign investment decision.

The paper is organized as follows. In Section 2 we derive a general definition of Business Groups on the basis of the existing literature and introduce our dataset, providing at the same time some stylized facts. In Section 3 we construct our metrics of vertical integration and organizational complexity and describe their properties. Section 4 relates our metrics of group boundaries to the home and host countries institutions in which BGs operate, as well as to the performance of affiliate firms within groups. Section 5 presents further lines of research and concludes.

2 The nature of Business Groups

2.1 Definition of Business Group

A commonly accepted definition of Business Groups does not exist in the economic or business literature, with Williamson (1975) already hinting at the fact that BGs should be located somewhere between markets and hierarchies.[8]

[8]Business and sociological studies also pointed out the difficulty to classify network-like forms of organization through a simple dichotomy of markets and hierarchies (see for example Powell, 1990; Granovetter, 1995; Hennart, 1991, 1993).

In their survey article, Khanna and Yafeh (2007) consider Business Groups as operating in multiple and often unrelated markets, but observe that they are formed by clusters of legally distinct firms with a common management, a characteristic that makes them different from multidivisional forms of organization. The finance literature emphasizes the groups' pyramidal structure built by a controlling shareholder through a chain of equity ties, and the possible conflicts of interests arising with minority shareholders (La Porta et al, 1999; Almeida and Wolfensohn, 2006). The focus of the industrial organization literature is instead on the creation of production chains through vertical integration within and across industries (see for example the survey by Lafontaine and Slade, 2007) or, in the case of international trade, through offshoring to foreign countries (among others Antràs, 2003; Grossman and Helpman, 2004). The phenomenon of BGs has also been extensively explored by the business literature, with a variety of different definitions summarized by Colpan and Hikino (2010). Briefly, scholars from different fields usually emphasize some attributes mostly related to their own field of study, generally ending with a working definition suitable for their research scope.

In this paper we argue that the lowest common denominator of all existing approaches is rooted in the nature of Business Groups as hybrid organizations of economic activities, halfway between markets and hierarchies. As such, BGs are thus able to exchange intermediate goods and services on the market, but possibly through a transfer price; they can relocate financial resources across affiliates, but at more favorable conditions if confronted with external financing, via the development of internal capital markets; they coordinate management decisions through majority stakes in controlled assets, but have to consider as well minority shareholders' protection. More generally, they have a flexible form of assets' ownership that provides at the same time incentives to self-inforce promises of cooperation among affiliates, given the control exerted by a common parent, without giving up the advantage (if and when necessary) of organizing activities within a market-like environment, since each affiliate maintains formal property rights on its production assets.[9]

We can thus define a Business Group as a set of at least two legally autonomous firms whose economic activity is coordinated through some form of hierarchical control via equity stakes. Legal autonomy and hierarchy are jointly constituent attributes of BGs, distinguishing them from independent firms (as these are legally autonomous but operate without hierarchies) and from multidivisional firms (which are organized through internal hierarchies of branches, but without autonomous legal status).[10]

Given the requirement of hierarchical control, our definition rules out strategic business alliances but includes in principle joint ventures, since their assets are owned (and controlled) by more than one proprietary firm. Under this general definition, multinational enterprises (MNEs) can also be considered as a special case of Business Groups, since they have by definition at least one legally autonomous affiliate located abroad, ultimately controlled by a parent located in the origin country.

[9]To this end, Baker, Gibbons and Murphy (2001a, 2001b) introduce the notion of relational contract: the decision to integrate or not is seen as dynamic in nature, as a repeated game, subordinate to the establishment of the parties' long term relationship. From this perspective, the emergence of Business Groups can be seen as a way to establish a superior relational contract, which facilitates integration or non-integration whenever needed.

[10]The notions of branches/divisions and subsidiaries/affiliates tend to overlap in some contexts. In this paper, in accordance with international standards (for example UNCTAD, 2009) we define a branch as a new location, division, department or office that is set up by a corporation, yet still within the original company's legal boundary. We will alternatively use the term subsidiary or affiliate for a legally independent company controlled by a parent.

In the case instead of economic entities with more than one productive plant (multi-plant firms), if all plants are commanded by the same firm under a single legal status we consider them as branches of that firm, as plants have no form of control on the production assets. On the other hand, if a plant has autonomous legal status, we consider it as an autonomous firm, thus either independent or an affiliate to a Business Group.

Our definition is wide enough to include either very simple groups with two firms, a headquarter and one affiliate, or very complex groups with hundreds of domestic and/or foreign affiliates linked by hierarchical control. Hence, we rule out any *ad hoc* definition in terms of minimum number of affiliates or industries, as found in some management or business literature (e.g. Colpan and Hikino, 2010).

2.2 Data

Having defined a Business Group as a combination of firms with autonomous legal status under some form of hierarchical control, the main difficulty in constructing a dataset on BGs is related to the identification of the notion of control exerted by a parent on affiliates.

We opt here for a definition of control as established in international standards for multinational corporations (OECD 2005; UNCTAD, 2009; Eurostat, 2007), where control is assumed if (directly or indirectly, e.g. via another controlled affiliate) the parent exceeds the majority (50.01%) of voting rights of the affiliate and can thus be considered as the Ultimate Controlling Institution / Ultimate Beneficial Owner.[11] Such a notion of control is not exhaustive, as it leaves outside the boundaries of BGs affiliates *de facto* controlled through minority ownership (<50%) as a result of a more fragmented property, or peculiar forms of control derived by some form of market advantage (e.g. a monopsony), as well as particular forms of government regulations (e.g. 'golden shares'). Yet, it has some clear advantages. First, the majority (50.01%) of voting rights criterion creates a unique standard for both domestic and multinational Business Groups. Second, it allows to rule out cases of double (or triple) accounting of affiliates among different groups, thus generating a definition of the boundaries of a BG which is univocal (technically, each of our business groups is a closed set). Third, such an approximation of control allows for a straightforward comparison with official statistics, as it is commonly used in international standards on foreign affiliates (Eurostat or OECD FATS) and for international tax purposes (IAS, IFRS).[12]

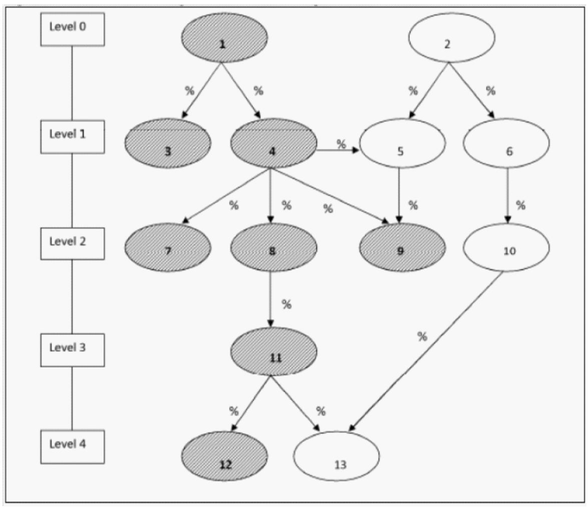
Figure 1 represents the organization of a typical Business Group as it can be derived by the application of the majority ownership notion of control. Such a representation corresponds to a mathematical object known as a hierarchical graph.[13] The upper shaded node (1) represents the

[11] Control derived by voting power, i.e. majority ownership, can be obtained through either direct or indirect cross-participations. A company X can control 60% of shares of company A, which controls 70% of shares of company B. Although company X does not formally control company B directly, it does indirectly, via company A. The latter, known as the principle of the Ultimate Controlling Institution in OECD FATS Statistics (or Ultimate Beneficial Owner in UNCTAD data), allows to assign control of company B to company X, thus called the parent company.

[12] A potential drawback of this methodology is that it can lead to an overestimation of control in some bigger networks of affiliates. See infra for a validation of our data and Vitali et al. (2011) for a reference on this issue.

[13] Technically, a hierarchical graph is a particular variation of a flat graph to which at least one parent node is added so as to assign functions to the other nodes (Palacz, 2003). Hierarchical graphs in turn are a generalization of a tree graph, in which several arms depart from one vertex as in a tree, but two different nodes are connected by only one edge; in hierarchical graphs, instead, different ultimate vertices can be directly or indirectly connected through several

Figure 1: Business Groups as hierarchical graphs



headquarter (or parent company), conventionally placed at level 0 of the hierarchy. The lower shaded nodes below level 0 represent the affiliates considered to be inside the boundaries of the same group, on their different hierarchical levels, with the edges connecting the nodes representing participation links.[14] The white nodes are instead firms possibly participated by the considered BG, but excluded from its boundaries on the basis of the majority ownership threshold.

Two different sorts of data have been combined to retrieve Business Groups: worldwide proprietary linkages provided by the Ownership Database by Bureau Van Dijk and firm-level financial accounts, from Orbis, by the same Bureau van Dijk.[15] Both proprietary linkages and financial data refer to the last available information available in year 2010. In Appendix A we provide a detailed description of our data sources and of the methodology we employ to identify Business Groups.

After considering (direct and indirect) control through majority ownership, we end up with 270,374 headquarters of Business Groups controlling a total of 1,519,588 affiliates in 207 countries in the year 2010. Given our hierarchical graph structure, firm-level data of affiliates are stratified according to their position in each Business Group, taking into account the level of proprietary distance from the headquarter. For each headquarter and each affiliate along the control chain we have industry affiliations at the 6-digit NAICS classification, including both primary and secondary activities, from which we can infer measures of vertical integration, as well as balance sheet data, from which we retrieve proxies of performance and productivity.

Not all firms in our dataset report a complete set of financial data. Moreover, country-level data for some institutional variables we use as controls are not available for every country. Hence, while

edges. Hence, different from a tree graph, in hierarchical graphs a parent node can coordinate other nodes at different hierarchical levels. Such a property makes them particularly suitable to visualize complex organization patterns such as the one represented by a BG.

[14]In this graph we interpret edges as control participations, but in a generic hierarchy of firms they could also represent trade flows of intermediate goods and services, or information flows for coordinated management actions.

[15]Other recent studies, including Acemoglu, Johnson and Mitton (2009) or Alfaro, Conconi, Fadinger and Newman (2010), exploit data sourced by Dun & Bradstreet (D&B). The latter is one of the sources now integrated in the Ownership Database by Bureau Van Dijk. For further details on the original data sources, see Appendix A.

we discuss here the complete dataset to introduce stylized facts on Business Groups, in our empirical strategies we rely on a restricted sample of data in which both firm-level and country-level information are available. The restricted dataset still encompasses 208,181 headquarters (groups) controlling a total of 1,005,381 affiliates in some 129 countries. The general properties described here also hold for the restricted sample of Business Groups.

In Table 1 we provide a geographical coverage of the whole sample by some main countries/areas. The headquarters of Business Groups (parents) are classified by their home country in the second column, while in the third column we report the total number of affiliates they control worldwide, either domestically or abroad, a distinction provided respectively in column 4 (domestic affiliates) and 5 (affiliates abroad, i.e. outward FDI by parents). In the last column we report the foreign affiliates located in the economy as the result of an inward FDI activity from parents abroad. Two thirds of Business Groups are originated in OECD economies, and those groups own around 76% of affiliates worldwide. The European Union, in particular, is in charge of 48% of affiliates, of which one third is located abroad. More than 50% of affiliates are located outside the home country in groups originating from OECD countries, especially in the US (85%), while the proportion is lower in developing countries (around 30%), where groups tend to be domestic.

Confronting the last two columns of Table 1, we observe a positive difference between outward and inward FDI stock (as proxied by number of affiliates) in developed economies, in particular in the case of US and Japan, where the number of affiliates located abroad outnumbers respectively more than twofold and fourfold the number of foreign affiliates located in the economy. European Union members seem an exception, but in that case it is intra-EU FDI activities that makes the net position almost in balance. In developing countries the inward FDI stock of firms is almost twice as large as the outward one.

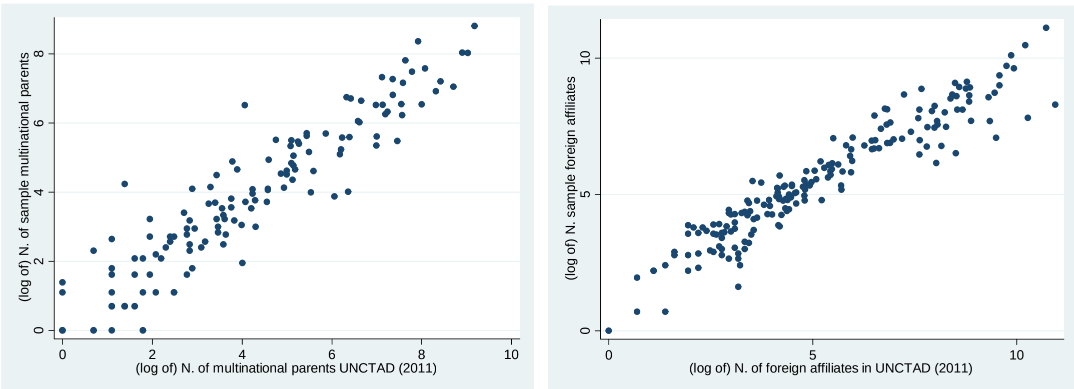
Table 1: Geographic coverage of Business Groups by headquarters and affiliates

Economy	N. of parents (Business Groups)	N. of affiliates (A + B)	Domestic affiliates (A)	Affiliates abroad (B)	Foreign affiliates located in economy
OECD	177,306	1,148,011	757,778	390,233	324,255
non-OECD	93,068	371,577	295,882	75,695	141,673
European Union	144,562	735,487	496,209	239,278	258,060
US	9,935	211,265	114,364	96,901	40,404
Rest of the world	115,877	572,836	421,441	151,395	167,464
of which:					
Japan	14,236	119,374	102,306	17,068	4,351
Latin America	3,972	11,480	7,106	4,374	18,656
Middle East	3,130	18,008	7,675	10,333	9,147
China	1,922	24,868	18,146	6,722	17,494
Africa	1,095	10,733	5,961	4,772	12,298
ASEAN	1,870	26,333	15,272	11,061	15,578
Total	***270,374***	***1,519,588***	***1,053,660***	***465,928***	***465,928***

To validate our dataset we can rely on few references since, to the best of our knowledge, there is no similar dataset covering control chains of corporate activities both domestically and abroad. One exception is the World Investment Report of UNCTAD, which compiles yearly a list of the biggest corporations currently operating in the world, all present in our dataset with their affiliates.

UNCTAD (2011) also reports the number of parents and affiliates involved in FDI activities hosted by each country. Based on these data, in Figure 2 we report the correlation between the number of headquarters controlling foreign affiliates abroad (left panel) and the number of foreign affiliates (right panel) located in each country as retrieved from our sample, against the similar figures provided by UNCTAD (2011): correlations are .94 and .93, respectively.[16]

Figure 2: Sample validation: (Logs of) numbers of multinational parents and foreign affiliates by host country in the sample and in UNCTAD (2011)



Finally, an indirect validation of the data is reported in Altomonte et al. (2012). In that paper, the authors have matched transaction- and firm-level data for France to the ownership structure of companies as derived from our dataset, in order to estimate the amount of intra-firm (intra-group) and arm's length (non intra-group) exports of French firms to the US in 2009. Looking at the counterfactual of official data on US intra-firm and arms' length imports from France, as retrieved from the US Census Bureau, the two trade flows match very closely.

2.3 Stylized facts on Business Groups

Table 2 shows how firms that are affiliated to Business Groups are on average bigger than non-affiliated firms along different dimensions (see Appendix A for information on the control group of non-affiliated firms): they employ on average 88% more workers, their sales are larger, they are usually more capital-intensive and almost twice more profitable. They are also 4% more productive, even after controlling for size and capital-intensity. Moreover, affiliation premia do not display dramatic differences between OECD and non-OECD economies.

In addition to the superior performance of BGs' affiliates, another typical characteristic found in the literature on heterogeneous firms is the remarkable skewness of the underlying distributions. In terms of hierarchies, the left panel of Figure 3 shows that 57% of firms in our dataset represent very

[16]The original source for data on affiliates in UNCTAD (2011) is Dun &Bradstreet, that is one of the sources of ownership data on which the ORBIS database also relies. The survey of UNCTAD (2011) refers to data in 2009, while our data are updated to 2010. We have excluded from the validation reported in Figure 2 the datapoint on China, since the country does not adopt the international standard definition of control (>50.01%) in reporting the number of affiliates, preferring a less committal criterion of 'foreign-funded enterprises', leading to non comparable figures.

Table 2: Premia for affiliates of Business Groups vs non-affiliated firms

Dependent variable	All countries	OECD economies	non-OECD economies
Log of employment	.88***	.90***	.80***
Log of turnover	1.32***	1.34***	1.15***
Log of capital	1.26***	1.25***	1.37***
Log of capital intensity	.30***	.29***	.35***
Log of profit	1.99***	2.01***	1.64***
Log of labor productivity(1)	.04***	.02**	.05***

Binary regressions with country-per-industry fixed effects; **, *** stand for significance respectively at 5% and 1%; (1) Capital-intensity added as a further control for a one-factor measure of productivity. See Appendix A for details on the control group of non-affiliates.

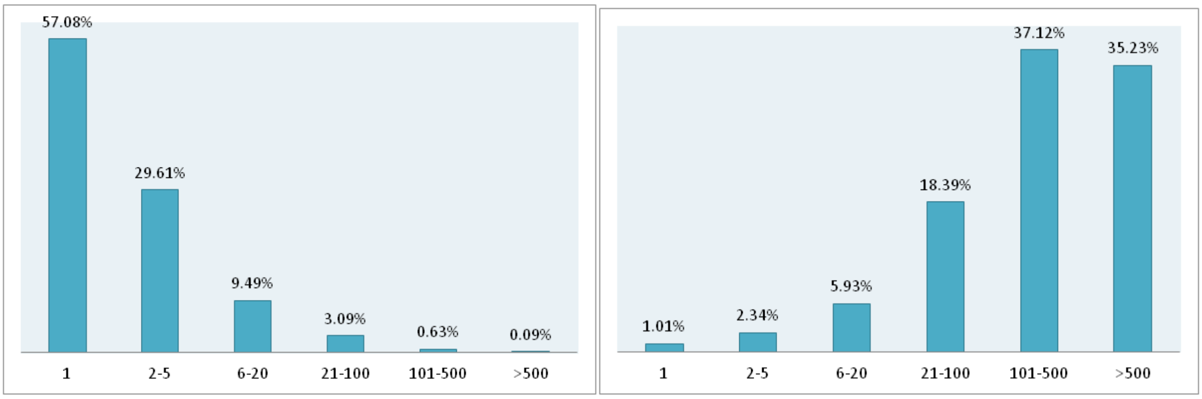
simple organizations consisting of one headquarter and one affiliate, while about 13% of groups have more than five affiliates and only 0.7% of headquarters control more than 100 affiliates. However, the right panel of Figure 3 also shows that those 0.7% of groups with more than 100 affiliates are responsible for more than 70% of value added recorded in our data.

The skewness in the distribution is in any case heterogeneous across countries, as shown in Table 3. US corporate groups tend to be larger, with an average size of 21 affiliates against a total average of 5, with largest groups operating in the financial industry and some in manufacturing. In Asian countries (Japan, China and the ASEAN region) we also detect the existence of conglomerates with a higher number of affiliates on each percentile of the distribution, as well as groups that tend to be internally engaged in all sectors of economic activities, from manufacturing to services.[17] In the case of Africa and Middle East, on the other hand, most of the bigger groups are active in the extraction of natural resources and related activities. European groups are on average smaller in terms of number of affiliates but there is a considerable difference between northern countries (Germany, Sweden, Finland, France) and southern countries (Italy and Spain), with the BGs originating from 'core' Europe being usually bigger than the ones originated in Southern Europe.

In the next Section we rely on the property rights theory of the firm and try to make sense of such a cross-country heterogeneity by linking some specific characteristics of Business Groups to the host countries' institutional environment.

[17] This is an inheritance of the former *keiretsu* or *chaebol* business groupings in countries like Japan or S. Korea, respectively.

Figure 3: Size distribution of Business Groups, number of affiliates vs value added



a) Overall distribution of affiliates of Business Groups (size classes)

b) Overall distribution of value added of Business Groups (size classes)

3 Metrics for Business Groups

3.1 Group vs. Affiliate Vertical Integration

Acemoglu, Jones and Mitton (2009) have explored in their paper the determinants of vertical integration in a large dataset of firms. They found that the contemporary presence of higher contracting costs and better financial development is associated to a higher firm-level vertical integration. That is, a single firm widens its boundary of economic activities in presence of both poor contract enforcement and good financial development, while contracting and financing constraints, individually considered, seem to have no effect on vertical integration.[18]

In absence of actual data on internal shipments of intermediate goods and services across firms, AJM (2009) proposed to proxy vertical integration exploiting the information on the set of industries in which a firm is engaged, combined with the input coefficient requirements that link those industries as retrieved from input-output tables. A firm-level index was therefore calculated summing up all input-output coefficients that linked each firm's primary activity to the secondary activities in which it was involved. The assumption is thus that a firm engaged in more industries, where backward and forward linkages in production are important, is supposed to have a higher capacity to source internally more inputs for its final output.[19]

In deriving these results, AJM(09) have however treated each firm in their sample as independent, that is neglecting the possibility that the degree of vertical integration can be a function of the coordinated management decision of a Business Group, where the decision to "make or buy" can be

[18]They also found that the impact of contractual frictions was more important in industries where holdup problems were more relevant. Hence, once industrial composition was accounted for, they concluded that some countries with a generalized problem of contractual incompleteness simply specialize in sectors where more vertical integration naturally occurs, that is in sectors where technologies are less advanced.

[19]For a previous attempt in the business literature, on which Acemoglu, Johnson and Mitton (2009) have built, see Fan and Lang (2000). For a similar application of this index see Alfaro et al. (2011).

Table 3: Descriptives of size distribution of affiliates by main countries/areas of origin

Home country	Mean	50 perc	75 perc	95 perc	99 perc	Max
OECD	6	1	3	17	94	2,707
non-OECD	4	1	2	13	46	996
European Union	5	1	3	13	65	2,557
US	21	3	9	92	354	2,707
Rest of the world	5	1	3	15	60	1,672
of which:						
Japan	8	1	4	31	119	1,000
Latin America	3	1	2	8	37	229
Middle East	6	1	4	19	69	492
China	13	3	9	40	127	574
Africa	10	2	9	42	116	455
ASEAN	14	5	13	50	155	479
Total	***5***	***1***	***3***	***16***	***74***	***2,707***

Descriptives reported by main countries/areas according to the origin of the parent company.

differentiated between headquarters and affiliates or across the same affiliates, as shown by the GM vs. Mitsubishi example.

To take into account the latter dimension, we have slightly refined the original AJM(09) index of vertical integration. First, we consider two layers of integration: the group-level, which is the result of all production activities performed by affiliates and headquarter altogether; and the affiliate-level, that is the propensity of each affiliate to exchange intermediates within the network represented by the group. Second, we take into account the number of lines of business in which a BG and its constituent firms can be involved.

In particular, we assume that within a group two sets of activities can be identified: a set of output activities $j \in N_H$, and a set of intermediate activities $i \in N_A$. The set of output activites coincides with the primary and secondary activities of the headquarter (N_H), whereas the range of intermediate activities at the group-level is represented by the set of primary and secondary activites in which controlled affiliates (N_A) are involved.

With these assumptions, we can build a group-specific input-output table as the one illustrated in Figure 4, where we report outputs in columns and inputs by row and where each combination VI_{ij} is the ith coefficient requirement to produce the jth output.

As in AJM(09) or Alfaro et al. (2011), we assume that industrial backward and forward linkages for all firms in our sample can be proxied by US input-output tables and adopt the industrial classification provided by the US Bureau of Economic Analysis, with 61 main industries mainly at a 3-digit level of disaggregation of the NAICS classification. In Appendix B we report the correspondence between the NAICS codes we retrieve from our data and the industries reported by the US Bureau of Economic Analysis.

By summing up input coefficient requirements by column in Figure 4 we obtain the vertical inte-

Figure 4: A group-specific input-output table

Outputs - headquarter's activities
$j \in N_H$

Inputs - affiliates' activities
$i \in N_A$

	j_1	j_2	j_3	...	j_{61}	Intermediate demand by input
i_1	VI_{i1j1}	VI_{i1j2}	VI_{i1j3}	...	VI_{i1j61}	*∑VI from input i_1*
i_2	VI_{i2j1}	VI_{i2j2}	VI_{i2j3}	...	VI_{i2j61}	*∑VI from input i_2*
i_3	VI_{i3j1}	VI_{i3j2}	$VI_{i_3j_3}$	...	VI_{i3j61}	*∑VI from input i_3*
...	...	...	...	...	...	...
i_{61}	VI_{i61j1}	VI_{i61j2}	VI_{i61j3}	...	VI_{i61j61}	*∑VI from input i_{61}*
Intermediate usage by output	*∑VI for output j_1*	*∑VI for output j_2*	*∑VI for output j_3*	...	*∑VI for output j_{61}*	

gration for each line of business in which the Business Group is involved.[20] To retrieve the vertical integration index for the whole group, we average the total of all input coefficient requirements (VI_{ij}) by the number of output activities ($|N_H|$), thus correcting for the potential conglomerate nature of the group.

The result is the following group-specific (g) vertical integration index:

$$v_g = \sum_{\substack{i \in N_A \\ j \in N_H}} \frac{1}{|N_H|} VI_{ij} \tag{1}$$

where VI_{ij} are the input coefficient requirements for any output activity $j \in N_H$ sourcing from all input activities $j \in N_A$. The group-specific vertical integration index can range from 0 to 1, where 1 corresponds to complete vertical integration.

The latter however does not capture the full picture of a BG's possible spectrum of choices in defining its boundaries. In fact, Business Groups could report similar levels of vertical integration at the level of headquarters, but they can organize each affiliate in a more or less integrated way, according to the organizational structure of the group across industries. The latter is the case of GM vs. Mitsubishi: as discussed, the former is a relatively specialized group, while the Japanese conglomerate is involved in more than ten lines of business. And yet, calculating an index of vertical integration at the level of headquarters as above (v_g) would yield similar results across the two groups. The reason is that affiliates in these two groups have themselves different degrees of vertical integration, which 'compensate' for the ex-ante different diversification of the headquarters' activities (affiliates of Mitsubishi tend to be bigger and present in more diversified sourcing industries then the ones of GM).

It then follows that estimating vertical integration in a sample that considers each BG's affiliate as an independent firm would clearly miss the structural correlation linking affiliates of the same group, thus generating potentially biased results. This is an important feature of Business Groups'

[20] As in AJM(09), in absence of actual data on internal shipments of intermediates, we can interpret this number as a mere propensity to be vertically integrated, where the sum of industry-level requirements gives us only the maximum possible integration of production processes.

boundaries which has been previously neglected in the analyses on vertical integration.

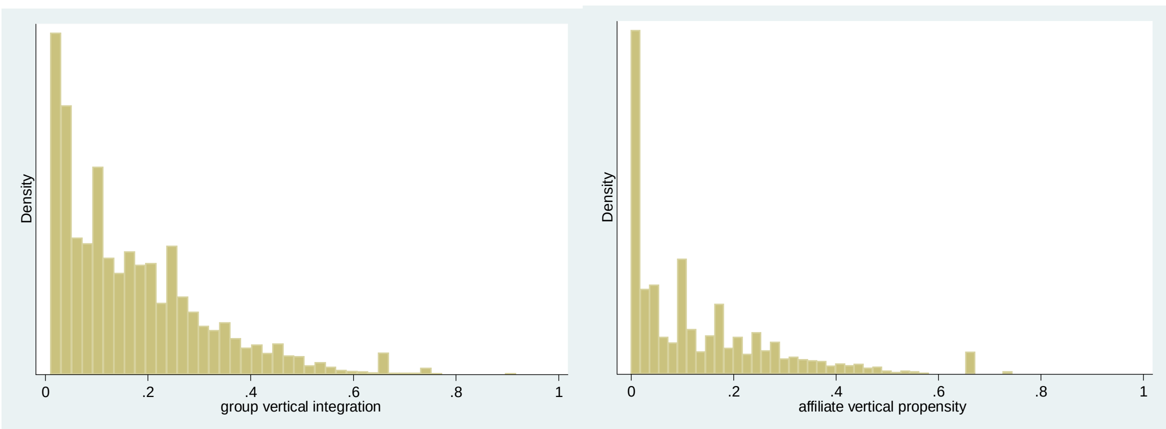
To better gauge the differences in vertical integration strategies across BGs, we thus integrate the group-index of vertical integration with a measure calculated directly at the individual affiliate level. Here we consider primary or secondary activities of the single affiliate as intermediate inputs that can be supplied potentially to all other co-affiliates and to the headquarter, and reclassify them according to the main industries reported in Appendix B. We end up with the following affiliate-specific (v_A) index of vertical integration:

$$v_A = \sum_{\substack{i \in N_f \\ j \in N_H}} \frac{1}{|N_f|} VI_{ij} \tag{2}$$

where the input coefficient requirements (VI_{ij}) are taken for any i among single affiliate activities ($N_f \subset N_A$) that can lend to any j main activity performed by the headquarter (N_H). Averaging by the number of main industries in which the single affiliate is involved allows again to correct for the potential conglomerate nature of the affiliate itself. In a nutshell, going back to Figure 4, this time we sum up coefficient requirements by row, then averaging by the number of rows. As well as for the previous group-specific index, the affiliate-level index can range from 0 to 1 and it can be interpreted as the propensity of an affiliate to be vertically integrated with the rest of the group.

Both the group- and affiliate-specific indices of vertical integration are additive on industries but not on production units: a new industry adds to the sum of input-output coefficients however small its contribution can be to the final output, but more firms can be involved in the same industry. For these reasons, we expect the group-level index of vertical integration to be higher than the same index calculated at the affiliate-level. In Figure 5 we report the sample distributions of both indexes.

Figure 5: Group-level and affiliate-level vertical propensities, sample distributions



a) Density calculated on a sample of 208,181 groups of firms. Mean: .062; standard deviation: .122; skewness: 2.723.

b) Density calculated on a sample of 1,005,381 affiliates; Mean: .049; standard deviation: .114; skewness: 3.189.

In our dataset the average vertical integration across groups (v_g) is .062 (that is, on average 6 cents worth of inputs are sourced within groups for a one dollar unit of output), while the same figure across individual affiliates (v_a) is .049. For comparison, the figure obtained by AJM(2009) on their (unconstrained) sample is of .0487, very similar to the one obtained in our data for the affiliate-level index. Alfaro et al. (2011) also calculated in a similar way a vertical integration index for manufacturing firms with more than 20 employees, obtaining an average vertical integration of .063 which is similar to the one we obtain for groups.

3.2 Hierarchical complexity

A particularly convenient property of representing Business Group as hierarchical graphs, as in Figure 1, is that it is possible to provide a synthetic measure of their complexity through some hierarchical form of entropy.

Borrowing from graph theory, it is possible to define a concept of entropy for a hierarchical graph G characterized by a total of L levels of hierarchies by assigning a discrete probability distribution $p : L \rightarrow [0,1]$ to every level l in the hierarchy, where the probability $p_l = \frac{n_l}{N}$ is a function of the n_l number of nodes on each level l and the total number of nodes N, yielding a measure of node entropy

$$H(G) = -\sum_{l} p_l \log(p_l) \tag{3}$$

which is specific for hierarchical graphs (Emmert-Streib and Dehmer, 2007).[21]

The $H(G)$ measure of entropy is characterized by some useful properties: a) it is continuous; b) it is additive in L, so that each level l (order) of nodes can be considered a subsystem of the whole graph G; c) the measure is maximal when all the outcomes are equally likely, i.e.there is an equal number of nodes on each level l. Finally, the logarithmic entropy is also symmetric, meaning that the measure is unchanged if levels L are re-ordered.

The symmetry of the measure is however an unpleasant property when applied to the case of Business Groups, since it implies that adding one node (affiliate) to the network increases its complexity independently of the hierarchical level at which the node is added, that is $\frac{\partial H(G)}{\partial p_m} = \frac{\partial H(G)}{\partial p_n}$ with $m \neq n$ being two different hierarchical levels. The latter is counter-intuitive in the case of a hierarchical organization characterized by a headquarter, because one might expect that the degree of coordination of the whole control chain (its 'complexity') should increase relatively more when affiliates are incorporated at proprietary levels more distant from the vertex.

For this reason we have refined the original $H(G)$ formula introducing an additional weight to the probability distribution of levels more distant from the parent. After some straightforward manipulations we can rewrite our node entropy measure for Business Groups, which we refer to as 'Group Index of Complexity' (GIC), as:

$$GIC = \sum_{l}^{L} l \frac{n_l}{N} \log\left(\frac{N}{n_l}\right) \tag{4}$$

where as before the measure is a function of the n_l number of affiliates on a given hierarchical level l, of the total number N of affiliates belonging to the group and of the total number of levels (L).

[21]Defining $p_l = \frac{n_l}{N}$ implicitly exploits a fundamental postulate in statistical mechanics or thermodynamics according to which the occupation of any state is assumed to be equally probable. Also note that this formula uses a base-2 logarithm, rather than the natural log, in order to obtain positive marginal complexity for $n_l > 1$.

The index can theoretically range within the $[0, +\infty)$ interval, with zero now indicating a very simple organization in which a headquarter controls one or more affiliates located just one level of control below ($l = 1$). Moreover, the index retains some desirable properties of the original node entropy, as it is (logarithmically) increasing in the number of hierarchical levels. We provide some detailed statistical properties of the GIC in Appendix C.

Importantly for our purposes, and contrary to the original hierarchical entropy measure $H(G)$, the GIC now allows to take into account the marginal increase in complexity brought about by affiliates added to lower hierarchical levels, since $\frac{\partial GIC}{\partial p_m} > \frac{\partial GIC}{\partial p_n}$ for $m < n$ (with $p_{n,m}$ being the usual probability measures defined above), provided that $n_l > l$. More specifically, the logarithmic weight assigned to the probability term $p = \frac{n_l}{N}$ of every level is such to increase the measure of complexity when more subsidiaries are included at *different* lower levels of distance, while the function is decreasing at the margin when affiliates are added at the *same* level.[22]

The economic rationale for a decreasing marginal complexity is associated to the idea that some economies of scale intervene when firms expand their network of affiliates horizontally, with new affiliates entering the network at the same hierarchical level, while coordination costs can become more and more important once the network enlarges and deepens, locating affiliates to further levels from the headquarter. This is in line with the literature on knowledge-based hierarchies (see for example Garicano, 2000, or more recently Caliendo and Rossi-Hansberg, 2012), according to which the optimal design of a management hierarchy is the result of a trade-off between knowledge and communication. A further layer of management increases the utilization of knowledge, for which some economies of scale are assumed, but at the same time it also increases the cost of communication along the hierarchy.

Accordingly, in our case hierarchical distance from the headquarter implies a higher 'fixed cost' of communication (hence our correction for node entropy in eq. 4), while further affiliates on the same level imply a decreasing 'marginal cost'. As a result, the hierarchical complexity of an object such as a Business Group cannot simply be proxied by its number of affiliates or by its number of hierarchical levels, and the index of complexity is not strictly monotonous in N. In Appendix C we provide further evidence of the sample comparison between groups' number of affiliates and our index of complexity.

Another way to measure the complexity of the hierarchy developed by a Business Group could be the explicit introduction of an edge entropy, i.e. considering the strength of the cross participations as a further dimension to be included in the entropy index. In this case, the index would differ if an affiliate can be finally owned through direct participation (held by the headquarter) or indirect cross participations (held by any other affiliates in the control chain).[23]

However, given the scope of our analysis, the latter would not yield qualitatively different results, as we only use data on business groups characterized by a majority threshold for control that includes

[22] This can be easily verified by taking the first derivative of Eq. 4 with respect to N or n_l. Note that now the maximum entropy is not reached when outcomes of states are equally likely (i.e. there is an equal number of affiliates at each level l). Rather, it is maximal when the group is pyramidal.

[23] In this case we could modify the index considering a joint probability distribution $p_{ij} = p_i^e * p_j^n$, such that $p_j^n = \frac{n_l}{N}$ as before and $p_i^e = \frac{e_l}{E}$ with e_l number of edges at level l and E total number of graph edges. The two events' probabilities can be assumed as mutually independent, and hence we obtain the following index $GIC^* = \sum_i^E \sum_J^L p_{ij} \log(1/p_{ij})$ where $\frac{\partial GIC^*}{\partial p_{ij}} < 0$, with $n_l, e_l \in \mathbb{N}$ and $n_l > 1$, $e_l > 1$, obtaining a decreasing marginal complexity in both nodes and edges, provided that we have at least one subsidiary and one control link on each level.

direct and indirect equity ties, in line with international business statistics. In terms of interpretation, that is equivalent to assume that, once the group boundaries are identified through control, any share above such a threshold would not significantly affect the complexity of the organization, as the headquarter would retain in any case the decision power.

4 Empirical results

4.1 Group boundaries and institutions

We begin our analysis by applying the empirical strategy developed by Acemoglu, Johnson and Mitton (2009) to our group-specific and affiliate-specific measures of vertical integration.

We assume here that the coordinated management of a Business Group decides the organization of production activities in two stages: first the group decides how much total vertical integration it wants to achieve.[24] Then, in a second stage, managers decide how to achieve the desired degree of vertical integration, distributing it between affiliates and headquarters and across affiliates, also based on the underlying hierarchical structure in which affiliates are placed.[25]

We thus test for the drivers of Business Groups' boundaries in two nested steps, first considering the drivers of group-level vertical integration, and then testing for vertical integration at the affiliate level given the choice of vertical integration at the group-level, further controlling for the level of hierarchical complexity. In the first specification, we take as a dependent variable the measure of group-level vertical integration (v_{gkc}) introduced in the previous section, which is specific for each group g located in country c and operating in a core industry k:

$$\begin{aligned} v_{gkc} = \beta_0 + \beta_1 X_{c_g} + \beta_2 Z_{c_g} + \beta_3 X_{c_g} Z_{c_g} + \beta_4 GIC_g + \beta_5 mne_g + \\ + \beta_6 \ln emp_g + \beta_7 \ln gdpc_{c_g} + \gamma_k + \varepsilon_{gkc_g} \end{aligned} \tag{5}$$

In this model, X_{c_g} and Z_{c_g} are the two proxies for country-level contract enforcement and financial development already employed in Acemoglu, Johnson and Mitton (2009). They are respectively the (opposite of) country-level average cost of a claim expressed as percentage of the total value of the claim[26] and the country-level ratio of private credit provided by all financing institutions to GDP.[27]

Three controls for the characteristics of business groups are included. The first is a proxy for the group size (employment, emp_g), obtained either directly from the heaquarter's balance sheet consolidated data, if available, or calculated summing up the employees of the headquarter and affiliates.

[24] At this stage we can assume that the group also decides where (at home or abroad) it wants to locate its activities, a decision which we take as exogenous in this paper.

[25] Although the latter is obviously a semplification of the coordination of managerial decisions within the group, we find support for this hypothesis in the finding by Atalay, Hortacsu and Syverson (2012), according to which acquired plants in US usually resemble the acquiring firms in terms of vertical integration. That is, they start shipping their production to locations that their acquirers had already been shipping to, and they produce outputs that their acquirers had already been manufacturing.

[26] The cost in court fees and attorney fees, where the use of attorneys is mandatory or common, expressed as a percentage of the debt value (World Bank, 2011a). The higher the cost the more difficult to enforce the contract. To ease interpretation of results, we have taken the opposite of this variable.

[27] Private credit by any financing institution to GDP for 129 countries, sourced by the work of Beck and Demirgüç-Kunt (2009), updated now regularly for the World Bank (2011b). This variable has been extensively used in some finance literature, see for example Rajan and Zingales (1998).

The second control is our entropy-like measure for hierarchical complexity (GIC_g), which controls for the fact that a higher level of vertical integration might be correlated to a more or less complex corporate structure. Finally, a binary variable (mne_g) controls if each Business Group owns affiliates operating outside from his home country.

As in Acemoglu, Johnson and Mitton (2009), we also control for the potential endogeneity of institutions to development, through the (log of) GDP per capita ($gdpc_{c_g}$) of the country where the headquarter is located, assumed to be the country of origin (home country) of the business group.

A set of 3-digit NAICS industry fixed effects (γ_k) is added to exclude that our results are the consequence of a peculiar industrial composition. On that, note that even though Business Groups can be active in more than one industry, we assign each group to the core 3-digit activity of their headquarters, that is the activity which we have used as output in our index of vertical integration where most of the value added is generated. Errors are clustered by country, and variables are standardized to obtain beta coefficients. Nested results are reported in Table 4.

Table 4: Group-level vertical integration, group complexity and institutional constraints

Dependent variable : Group vertical integration	I	II	III	IV
contract enforcement	-.139***		-.114***	-.116***
	(.037)		(.037)	(.037)
financial development		-.085***	-.070**	-.071**
		(.035)	(.027)	(.028)
contract enforcement*financial development			.023	.020
			(.024)	(.024)
group complexity				.073***
				(.024)
multinational				-.056
				(.035)
(log of) group employment	-.003	.001	.003	.003
	(.006)	(.006)	(.004)	(.003)
(log of) GDP per capita	-.234***	-.229***	-.188***	-.185***
	(.059)	(.079)	(.056)	(.057)
Constant	2.290***	2.247**	1.812***	1.838***
	(.633)	(.818)	(.582)	(.583)
3-digit industry fixed effects	Yes	Yes	Yes	Yes
Errors clustered by country	Yes	Yes	Yes	Yes
Observations (N. of Business Groups)	222,433	222,433	222,433	222,433
Industries	88	88	88	88
Countries	129	129	129	129
Adjusted R squared	.357	.361	.376	.377

*, **, *** significance at 10%, 5% and 1%. Beta coefficients and errors clustered by country.

Results show that contracting and financial conditions on a country-level are both significantly and separately correlated with a group's vertical integration, even after controlling for industrial composition. We find in particular that a better contract enforcement reduces the scope for vertical integration, since in this case Business Groups can rely on external suppliers for the provision of inputs with a lower probability that they renege on commitments. Similarly, our results also show that a higher level of financial development reduces the necessity to internalize production activities: as credit constraints are less stringent thanks to the availability of better capital markets, outsourcing outside the boundaries of the group is the preferred strategy.

These results are in line with the general priors of the literature and only slightly different from the ones presented by Acemoglu, Johnson and Mitton (2009): in their case firm-level vertical integration at the country-level was found to be positively correlated with the interaction term between contracting instutions and financial frictions (not significant in our case), while the individual variables in their estimates were correctly signed (as in our case) but individually not-significant. We believe this difference in results is due to our choice of explicitly considering group affiliation in the construction of the vertical integration index.[28]. It is not completely clear in existing literature how contractual and financial frictions combine together in determining the level of vertical integration. Here we presume the perspective adopted by Acemoglu, Antràs and Helpman (2007), who showed theoretically how they can be complementary in the choice of the organizational form between a firm and its suppliers (see also Hart and Moore, 1990, for a previous mention of this complementarity). According to them, a firm faces a credit constraint when investing in a contractual relationship with its suppliers, because it needs to pay upfront for that investment. But credit markets can be imperfect and in this case they demonstrate that even in presence of a low cost of contracting, important financial frictions can lead to a choice of vertical integration.

Finally, we also find that the level of total integration is not different for multinational and domestic BGs, as the control in the last column of Table 4 confirms, which further strenghtens the idea that the home country institutional environment is a powerful driver of the organization of a Business Group.

Given the ability of Business Groups to design vertical integration also across affiliates, we nest the above results in the vertical integration choice of each affiliate, by estimating the following equation:

$$\begin{aligned} v_{a(g)kc} = {} & \beta_0 + \beta_1 X_{c_a} + \beta_2 Z_{c_a} + \beta_3 X_{c_a} Z_{c_a} + \beta_4 GIC_g + \beta_5 GIC_g X_{c_a} + \\ & + \beta_6 GIC_g Z_{c_a} + \beta_7 v_g + \beta_8 v_g X_{c_a} + \beta_9 v_g Z_{c_a} + \beta_{10} \ln gdpc_{c_a} + \\ & + \beta_{11} \ln emp_a + \beta_{12} \ln emp_g + \gamma_k + \varepsilon_{a(g)kc_a} \end{aligned} \tag{6}$$

where in this case we take as dependent variable the affiliate-specific (a) vertical integration within the gth group ($v_{a(g)kc_a}$), defined in Equation (5) as the average propensity to ship intermediate inputs within the group network.

Each affiliate is characterized by a core activity (k), where we assume most of value added is created (even though the affiliate can be involved in more than one primary and/or secondary activities), and by a country (c_a) in which the affiliate is located, possibly different from the country of origin of the Business Group, in which case we will be dealing with a foreign affiliate.[29] Hence, the set of proxies of institutional frictions (X_{c_a}, Z_{c_a}), their interaction and the (log) of GDP per capita ($gdpc_{c_a}$) now all

[28]Indeed, in the robustness and sensitivity checks we present in Table 6, we report in Column 1 the results of the above exercise carried out exactly as in AJM(09), that is ignoring the property linkages among firms when constructing the vertical integration indexes. As in their case, we now also get correctly signed but poorly significant coefficients.

[29]In Table 6 we will specifically test for the robustness of our results when dealing with foreign affiliates only.

Table 5: Affiliate-level integration, group complexity and institutional constraints

Dependent variable : Affiliate integration within the group	I	II	III	IV	V	VI
contract enforcement	-.128***		-.104**	-.094**	-.095**	-.099**
	(.061)		(.045)	(.043)	(.044)	(.043)
financial development		-.151***	-.138***	-.136***	-.140***	-.142***
		(.040)	(.041)	(.039)	(.042)	(.038)
contract enforcement*financial development			.018	.019	.018	.007
			(.034)	(.033)	(.035)	(.031)
group integration				.079***	.075***	.068***
				(.020)	(.022)	(.015)
group integration*contract enforcement					-.037***	-.018**
					(.009)	(.008)
group integration*financial development					-.080***	-.070***
					(.012)	(.006)
group complexity						-.012
						(.014)
group complexity*contract enforcement						.019
						(.015)
group complexity*financial development						.035***
						(.011)
(log of) group employment				-.002	-.002	-.002
				(.004)	(.004)	(.003)
(log of) affiliate employment	-.019	-.009	-.013	-.013	-.011	-.011
	(.012)	(.009)	(.009)	(.008)	(.009)	(.008)
(log of) GDP per capita	-.289***	-.163	-.146*	-.133*	-.148*	-.105
	(.094)	(.100)	(.079)	(.071)	(.084)	(.069)
Constant	2.980***	1.671	1.502*	1.374*	1.384*	1.070
	(.994)	(1.058)	(.824)	(.749)	(.751)	(.726)
3-digit industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Errors clustered by country	Yes	Yes	Yes	Yes	Yes	Yes
Observations (N. of affiliates)	1,005,381	1,005,381	1,005,381	1,005,381	1,005,381	1,005,381
Industries	110	110	110	110	110	110
Countries	129	129	129	129	129	129
Adjusted R squared	.400	.402	.410	.415	.417	.421

*, **, *** significance at 10%, 5% and 1%. Beta coefficients and errors clustered by country.

refer to the affiliate hosting country.

The inclusion of both the group index of complexity (GIC_g) and the group vertical integration (v_g) as covariates is crucial in our setting to comprehend how business groups solve the trade-off between vertical integration at the headquarter vs. affiliate-level within a given institutional setting. For this, we will also interact these two covariates with our proxies of institutional frictions.

Group-level and affiliate-level employment (emp_a, emp_g) in logs are also added as controls, as well as a set of NAICS 3-digit industry fixed effects (γ_k) that take into account potential differences in the industrial composition of the sample, while errors are clustered by country. Results are reported in Table 5.

Similar to the results of the group-level specification, we observe that the affiliate-level vertical integration is negatively correlated with contract enforcement and financial development: better institutions not only reduce the scope for total vertical integration, but also allow single affiliates to be relatively more specialized in the production of fewer inputs required by common production processes.

We also find that, on average, affiliate and group integration are positively correlated, as similar environments lead the single affiliate to resemble the parent group when designing its boundary, in line with the findings of Atalay et al. (2012).

The negative signs in the interaction terms between group integration and institutional quality measures (column 5) point however to the idea that the better the institutions in a country, the higher is the flexibility of a group in selecting the degree of specialization of its affiliates. In other words, in 'good' institutional environments affiliates are relatively less similar in terms of vertical integration with respect to their group: a higher contract enforcement and/or a better financial development allow the single affiliate to specialize more with respect to the group to which it belongs, thus exchanging fewer inputs within the same group.

Finally, in column 6 of Table 5, we introduce as a further control our new metric of organizational complexity and its interaction with institutions. Results point at a potential trade-off in the organizational design of a Business Group's boundaries between affiliate integration and complexity: controlling for the level of financial development, the group can choose either to increase the degree of vertical integration of each affiliate in the group, and in turn reduce the hierarchical complexity of the network; or it can rely on more 'specialized' affiliates, in this case organized within a more hierarchical structure. Notably these results are in line with the case studies of GM and Mitsubishi, both showing similar levels of group integration but a different degree of affiliates' integration (less integrated in the case of GM) vs. organizational complexity (with GM's affiliates being placed in a relatively more complex hierarchical structure).[30]

The latter findings confirms the idea that within Business Groups vertical integration choices are not independent from decisions on the organization of the hierarchy of firms across groups. Ignoring this latent organizational variable when checking for the drivers of vertical integration of firms, as well as the correlation between group and affiliate level of vertical integration (e.g. considering all firms as independent in a sample) might thus lead to an unobserved variable bias.

In Table 6 we present some tests of sensitivity of our results. In the first column, as already discussed, we reproduce for our sample the original methodology by Acemoglu, Johnson and Mitton (2009). That is, we calculate the index of vertical integration for each firm as in AJM(09), i.e. assuming that each firm's primary activities are its outputs while its secondary activities are internally produced inputs, but ignoring the property linkages among firms when constructing the vertical integration measures, thus considering each firm as independent. Results show that, as expected, when omitting the Business Group dimension the model is less well specified, and thus institutional variables lose significance.

In the second column of Table 6 we restrict our sample only to affiliates with more than 20 employees, in order to check if previous results are driven by the presence of a larger set of small firms, in which vertical integration can be assumed to be negligible. Coefficients remain very similar

[30] Clearly both vertical integration and organizational complexity can be endogenous to institutions: in an exercise not reported here, we have tested the correlation between BG's organizational complexity and institutional frictions (controlling for industrial composition and firm-level characteristics), finding that the relationship is positive and significant only in the case of financial development. However, to the best of our knowledge, there is no previous contribution investigating the direction of causality between institutions, vertical integration and organizational complexity, nor we have time-varying measures of the latter that allow for such a test. As such, the evidence reported here should be interpreted as correlations.

Table 6: Sensitivity of results: affiliate-level vertical integration

Dependent variable : Affiliate integration	AJM_09	> 20 employees	> 1 affiliate	domestic affiliates	foreign affiliates
contract enforcement	-.042*	-.104**	-.094**	-.106**	-.054*
	(.024)	(.041)	(.042)	(.041)	(.026)
financial development	-.012	-.151***	-.149***	-.152***	-.101**
	(.025)	(.037)	(.037)	(.038)	(.046)
contract enforcement*financial development	.003	-.003	.003	.024	-.022
	(.017)	(.028)	(.030)	(.031)	(.032)
group integration		.068***	.061***	.075***	.030***
		(.015)	(.013)	(.025)	(.008)
group complexity		-.012	-.016	-.008	-.014*
		(.014)	(.014)	(.014)	(.006)
group integration*contract enforcement		-.020**	-.015*	-.013	-.006*
		(.008)	(.008)	(.012)	(.002)
group integration*financial development		-.064***	-.064***	-.081***	-.028***
		(.006)	(.006)	(.011)	(.006)
group complexity*contract enforcement		.018	.015	-.005	.004
		(.015)	(.015)	(.015)	(.006)
group complexity*financial development		.033***	.038***	.015**	.040*
		(.011)	(.011)	(.007)	(.018)
(log of) group employment		-.001	-.002	.005	-.001
		(.004)	(.003)	(.004)	(.001)
(log of) affiliate employment	-.003	-.012	-.010	-.012	-.003
	(.004)	(.008)	(.008)	(.009)	(.007)
(log of) GDP per capita	-.086**	-.088	-.096	-.130**	-.013
	(.037)	(.068)	(.069)	(.064)	(.070)
Constant	.872**	.908	.979	1.321*	.046
	(.376)	(.715)	(.725)	(.672)	(.074)
3-digit industry fixed effects	Yes	Yes	Yes	Yes	Yes
Errors clustered by country	Yes	Yes	Yes	Yes	Yes
Observations (N. of affiliates)	1,005,381	831,319	875,727	748,453	256,928
Industries	110	110	110	108	109
Countries	129	129	129	129	129
Adjusted R squared	.215	.445	.422	.469	.291

*, **, *** significance at 10%, 5% and 1%. Beta coefficients and errors clustered by country.

to previous results also in magnitude. In the third column we exclude the simplest Business Groups characterized by only one affiliate and one headquarter, to verify the extent to which results are driven by the relatively large presence of groups of this kind in our sample: results are quite stable and all previous comments can be considered valid.

Findings are also robust (albeit slightly less significant) also when considering separately foreign affiliates, as shown in the fourth column of Table 6. Part of this loss of significance can be explained by the smaller variance we notice in the subsample of foreign affiliates, since they are more similar than domestic affiliates in terms of vertical integration. On the other hand we can assume that, with respect to the average affiliate in our sample, affiliates of multinational corporations are more influenced by organizational strategies developed in their country of origin, and hence relatively less influenced in their organizational design by host country characteristics.

4.2 Group boundaries and performance

Another interesting dimension in the analysis of Business Groups is the relationship between vertical integration and productivity. Atalay, Hortacsu and Syverson (2012) have empirically investigated US plant-level data showing that vertically integrated plants have on average higher productivity levels. They also find that this productivity premium reflected a cherry-picking effect, with already more efficient plants integrated *ex post* into more vertical structures.[31]

Consistently with these findings, in Table 2 we have shown that also in our dataset firms that are affiliated to Business Groups (and thus to some extent vertically integrated) are on average bigger and more productive than non-affiliated firms. What remains to be seen, however, is the extent to which different levels of vertical integration map into a different productivity of affiliates, considering that vertical integration for Business Groups is a multi-dimensional concept (at the headquarter vs. affiliates' level, and across affiliates distributed along more or less complex hierarchies). To explore these issues, we will thus test if the productivity levels of BGs' affiliates are systematically correlated to vertical integration and/or hierarchical complexity of the group, controlling for a number of additional groups' characteristics.

In our specification we take as dependent variable (the log of) labor productivity ($prod_{a(g)kc_a}$) calculated as value added per employee of each affiliate a belonging to the gth group, operating in core industry k and located in country c_a.

$$\ln prod_{a(g)kc_a} = \beta_0 + \beta_1 v_{a(g)kc_a} + \beta_2 v_g + \beta_3 GIC_g^{\theta} + \beta_4 hdist_{a(g)kc_a} + \beta_5 for_{a(g)kc_a} + \\ + \beta_5 \ln emp_{a(g)kc_a} + \beta_6 \ln emp_g + \beta_7 \ln kl_{a(g)kc_a} + \lambda_{kc_a} + \varepsilon_{a(g)kc_a} \quad (7)$$

Affiliate-level and group-level vertical integration ($v_{a(g)kc_a}$, v_g), as well as hierarchical complexity (GIC_g^{θ}) with $\theta = 1, 2$ are included as covariates, together with the hierarchical distance ($hdist_{a(g)kc_a}$) of each affiliate a within group g. The latter is the level at which the single firm is located within the network of affiliates that form a Business Group, as depicted in Figure 1, and can be interpreted as a control for the communication ability of the affiliate with the center of decision represented by the headquarter. Controls for capital intensity and size at the affiliate level ($emp_{a(j)kc_a}$, $kl_{a(j)kc_a}$) correct for the possible bias deriving from the use of a one-factor productivity indicator, at the same time controlling for relation-specific investments that a firm with a higher capital-intensive production can undertake. Total employment is included as a control at the level of the group (emp_g), together with a full set of (country-*per*-industry) fixed effects (λ_{kc}), in order to neutralize at this stage of the analysis all possible differences in institutional environments combined with industrial composition (here considered at the 3 digit level of disaggregation), and thus isolate as much as possible the effects of organizational design on affiliates' performance. Errors are clustered at the headquarter level, to account for within-group correlation. Nested results are presented in Table 7.

Looking at results, when we do not control for country fixed effects, as in the first column of

[31] For other works showing a positive correlation between vertical integration and productivity, see for example Maksimovic and Phillips (2002) or Schoar (2002), both investigating conglomerate firms. In both cases a selection effect is claimed to be responsible for the productivity differences between integrated and non-integrated firms.

Table 7, we obtain a negative correlation between vertical integration and productivity, both at the group- and at the affiliate-level. This is because Business Groups and their constituent firms are more vertically integrated in developing economies, where institutional frictions are more present and firm performance is on average lower than in developed economies. On the other hand, including country fixed effects, but excluding industry fixed effects (column 2 of Table 7), we find a positive correlation between both indexes of vertical integration and productivity, although in this case several omitted variables can bias the correlation, among which the degree of market competition and the specific contractual completeness of the industry in which the firms operate. This is why starting from column 3 we include country-per-industry fixed effects.

Table 7: Productivity levels and Business Groups' dimensions

Dependent variable :	OLS fe	OLS fe	OLS fe	OLS fe	OLS fe	OLS fe
log of (labor productivity)	(I)	(II)	(III)	(IV)	(V)	(VI)
affiliate integration	-2.092***	.377***	.133**	.106*	.104*	.104*
	(.085)	(.055)	(.059)	(.063)	(.063)	(.063)
group integration	-.587***	.239***	.055	.041	.035	.034
	(.097)	(.050)	(.037)	(.039)	(.040)	(.041)
group complexity	.062***	.020***		.009***	.038***	.041***
	(.004)	(.003)		(.002)	(.005)	(.005)
group complexity^2					-.002***	-.002***
					(.000)	(.000)
foreign			.283***	.256***	.246***	.250***
			(.012)	(.012)	(.012)	(.012)
hierarchical distance						-.012***
						(.004)
(log of) group employment	.073***	.027***	.019***	.015***	.013***	.013***
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
(log of) affiliate employment	.264***	-.037***	-.030***	-.020***	-.022***	-.022***
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
(log of) capital intensity	.264***	.171***	.181***	.181***	.180***	.180***
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
Constant	3.956***	4.337***	4.294***	4.295***	4.280***	4.295***
	(.020)	(.016)	(.017)	(.015)	(.015)	(.015)
3-digit industry fixed effects	Yes	No	No	No	No	No
Country fixed effects	No	Yes	No	No	No	No
Country*industry fixed effects	No	No	Yes	Yes	Yes	Yes
Errors clustered by headquarter	Yes	Yes	Yes	Yes	Yes	Yes
Observations (N. of affiliates)	219,368	219,368	219,368	219,368	219,368	219,368
N. of Business Groups	64,026	64,026	64,026	64,026	64,026	64,026
Industries	105	105	105	105	105	105
Countries	129	129	129	129	129	129
adjusted R_squared	.288	.164	.479	.487	.488	.488

*, **, *** significance at 10%, 5% and 1%. Errors clustered by headquarter.

Controlling for country-level heterogeneity combined with industrial composition (column 3), we find that only affiliate vertical integration is associated to average affiliates' productivity, while group integration is not significant. The latter result is confirmed also controlling for foreign affiliates (which in turn, consistently with other findings in the literature, are found to be some 25% more productive than the average firm). However, when also controlling for the hierarchical complexity of the group (column 4), we find that all the measures of vertical integration lose both significance and magnitude, while group complexity appears to be positively and significantly related to productivity.

This result is partially in line with the evidence provided by Atalay, Hortacsu and Syverson (2012) in the case of US data, according to which much of the positive correlation between plant performance and vertical ownership structures fades away when controlling for firm size as proxied by total revenues, employment or number of establishments. In our strategy we can indeed distinguish between actual (affiliate or group) size, measured by (affiliate or group) employment in the above specifications, and hierarchical complexity, being able to show that only after introducing this latter dimension the correlation between affiliates' productivity and vertical propensity fades away.[32]

Similarly to Atalay, Hortacsu and Syverson (2012), we believe that the main economic rationale that can explain at the same time a weak correlation between firm performance and vertical integration as well as a stronger positive correlation between productivity and organizational complexity is to be found in the theoretical framework developed by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007). According to these works, intangible assets (such as best managers, best managerial practices) are complementary to physical inputs involved in vertically linked products. Therefore, given a firm size (in our case group size), best intangible assets can be shared in presence of a larger number of units of production (in our case more complex hierarchies) and hence their costs can be smoothed on a larger scale.[33]

In light of this theoretical framework, we assume that gains of productivity can derive both from the exchange of intermediate inputs and from the exchange of intangible assets, with the latter channel of transmission showing a higher significance from our data. The relevance of intangible assets is also confirmed by looking at column 5 of Table 7, in which we include a squared term in hierarchical complexity that appears with a negative and significant coefficient. The latter can be attributed to coordination costs that arise when business networks become too complex to be managed efficiently. This result is also in line with the microfoundation of organization provided by Caliendo and Rossi-Hansberg (2012), in which a minimum efficient scale exists in the acquisition and communication of knowledge throughout the hierarchy, associated to the emergence of endogenous communication costs of additional management layers.

Although we do not have information on a group's minimum efficient scale of production, from our estimates we can calculate the optimal threshold of complexity after which, ceteris paribus, returns from hierarchical complexity start to decrease: this is quite large, as it corresponds to a GIC of around 9.5, associated to groups exceeding the number of 550 affiliates and/or organized in control chains with over 5 levels of hierarchical distance.[34] Such an evidence of marginally decreasing returns from

[32]In other words, in our data we are able to explicitely measure proprietary linkages among firms and their hierarchies, while Atalay, Hortacsu and Syverson (2012) limited their analysis to plants within firms, without an explicit control for ownership structures.

[33]Alternatively, we cannot exclude that the positive correlation between organizational complexity and affiliates' performance is driven by a higher degree of delegation of power as in Bloom, Sadun and Van Reenen (2012), according to which more delegation within the firm is associated with a better allocation of resources. The latter would imply that a more complex hierarchy in the chain of control, which we measure, is associated to a larger degree of managerial delegation from the headquarters to the affiliates, an assumption which we cannot however test, as internal delegation is unobservable to us.

[34]Note that our Group Index of Complexity is not strictly monotonous in mapping organizational forms, and thus a

increasing complexity is relevant, as it puts a natural limit to the growth in complexity of business groups: indeed, only 1% of groups in our sample (the critical value of GIC is around the 99th percentile of its distribution) exceed this average 'optimal' organizational threshold.

The same theoretical framework of Caliendo and Rossi-Hansberg (2012) can also explain another result we have obtained when distinguishing between hierarchical complexity (which takes into account the overall density of affiliates at each level of the control hierarchy) and the simple hierarchical distance, i.e. the length of the command chain linking each affiliate to the parent company. In general, we find that the further the firm is from the decision making center, the lower its level of productivity appears to be; however, the latter result only holds when we control for our main measure of organizational complexity. When considering only hierarchical distance in the model (itself a raw proxy of organizational complexity), affiliates located at further levels of control would actually display higher levels of productivity. Our data are thus consistent with the idea that further layers of management allow for the exploitation of economies of scale for knowledge inputs, and hence affiliates to bigger (more complex) networks are relatively more productive. At the same time, subsidiaries located at further hierarchical distances from the headquarter discount a higher cost of communication and show (at the margin) a negative productivity premium.

5 Conclusions

In this paper we have provided a preliminary comprehensive picture of the phenomenon of Business Groups across the globe, showing how BGs, although more numerous in developing economies, constitute however a relevant share of economic activities also in developed countries.

By refining traditional measures of vertical integration, we have shown how the decision process related to the design of boundaries in Business Groups is truly multi-dimensional, as vertical integration choices are not independent from decisions on the organization of the hierarchy of firms across groups, another feature of the organization of Business Groups for which we provide a novel metric. Ignoring the correlations among these variables when checking for the drivers of vertical integration of firms (e.g. considering all firms as independent in a sample, as the literature has been doing insofar) might thus lead to an unobserved omitted variable bias.

Another interesting result of the analysis points at a positive and significant relationship between organizational complexity and affiliates' productivity (controlling for vertical integration), a finding consistent with the theoretical framework by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007), according to which intangible assets are complementary to physical inputs involved in vertically linked products and a higher scale of production allows to smooth the costs of these best inputs.

The relevance of intangible assets is also confirmed by the non-linear relation between complexity and productivity that we find in the data, attributable to a trade-off between the exploitation of knowledge as an input and the communication of it down the hierarchy, a result further in line with the work by Caliendo and Rossi-Hansberg (2012) and in general with previous literature on knowledge-based hierarchies

given value can identify a range of cases with different combinations in terms of numbers of affiliates or control levels.

On the other hand we cannot exclude (but cannot test) an interpretation of this result in the light of the findings by Bloom, Sadun and Van Reenen (2012), according to which more delegation of power within the firm is associated with a better allocation of resources. Clearly, here the delegation of power has to be considered within the group, with additional layers of control across affiliates (and thus higher complexity) acting as a proxy for delegation.

A number of further lines of research stem from the above analysis. First of all, it is unclear if the correlation between organization and productivity is the result of a cherry-picking process, in which bigger and/or more complex business groups select firms with the better prospects on the market. To recover some evidence on the latter direction of causality we have used a subset of our data for which we have information on the date of acquisition of a firm by a Business Group. We have found that the relation holds also using the growth rates of productivity, but due to the persistency over time of the organizational complexity variable, more work needs to be undertaken in this direction.

A second line of research should investigate the country and industry variation in BGs' foreign affiliates to verify if and how a differential in institutional constraints between origin and host countries can shape organizational designs and finally affect performance.

References

Acemoglu D., Antràs P., Elhanan H. (2007). "Contracts and Technology Adoption", American Economic Review, American Economic Association 97(3), 916-943.

Acemoglu, D., Johnson S., and Mitton T. (2009). "Determinants of Vertical Integration: Financial Development and Contracting Costs". The Journal of Finance LXIV (3), pp. 1251-1290

Acharya V. V., Myers S. C. and Rajan R. G. (2011). "The Internal Governance of Firms", Journal of Finance 66(3), 689-720.

Aghion, P.and Holden R. (2011). "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?" Journal of Economic Perspectives, vol. 25(2), pp. 181–97

Alfaro L., and Charlton A. (2009). "Intra-industry Foreign Direct Investment", American Economic Review, 99 pp. 2096-2119.

Alfaro L., Conconi P., Fadinger H. and Newman A. F.(2011). "Trade Policy and Firm Boundaries", NBER Working Paper N. 16118, National Bureau of Economic Research.

Almeida H. V. and Wolfenzon D. (2006) "A Theory of Pyramidal Ownership and Family Business Groups", The Journal of Finance 61(6), pp. 2637-2680.

Altomonte C., Di Mauro F., Ottaviano G. I. P., Rungi A. and Vicard V. (2012). "Global Value Chains during the Great Trade Collapse: a Bullwhip Effect?", ECB Working Paper N. 1412, European Central Bank.

Antràs, P. (2003). "Firms, Contracts, and Trade Structure", The Quarterly Journal of Economics 118(4), pp. 1375-1418.

Antràs P. and Helpman E. (2004). "Global Sourcing," Journal of Political Economy, vol. 112(3), pp. 552–580.

Antràs P. and Helpman E. (2009) "Contractual Frictions and Global Sourcing," in E. Helpman, D. Marin, and T. Verdier (eds.), The Organization of Firms in a Global Economy, 2008, Harvard University Press, pp. 9-54.

Antràs, P. and Rossi-Hansberg E. (2009). "Organizations and Trade," Annual Review of Economics, Annual Reviews, vol. 1(1), pp. 43-64.

Atalay E., Hortacsu A. and Syverson C.(2012). "Why Do Firms Own Production Chains?", NBER Working Paper N. 18020, National Bureau of Economic Research.

Baker, G., Gibbons, R. and Murphy, K. J. (2001a). "Bringing the Market Inside the Firm?", American Economic Review, American Economic Association, 91, pp. 212-218.

Baker, G., Gibbons, R. and Murphy, K. J. (2001b). "Relational Contracts and the Theory of the Firm", The Quarterly Journal of Economics 117 (1), pp. 39-84

Beck T. and Demirgüç-Kunt A. (2009), "Financial Institutions and Markets Across Countries and over Time: Data and Analysis", World Bank Policy Research Working Paper No. 4943.

Beirlant J., Vynckier P. and Teugels J. (1996). "Excess Functions and Estimation of Extreme-Value Index". Bernoulli vol. 2(4), pp. 293-318

Bloom N., Sadun R. and Van Reenen J. (2012). "The Organization of Firms across Countries", The Quarterly Journal of Economics, 127(4), pp. 1663-1705.

Bloom N. and Van Reenen J. (2010). "Why Do Management Practices Differ across Firms and Countries?", Journal of Economic Perspectives vol 24(1), pp. 203–224.

Caliendo L. and Rossi-Hansberg (2012). "The Impact of Trade on Organization and Productivity", The Quarterly Journal of Economics 127(3), pp. 1393-1467.

Coles, S. (2001). An Introduction to Statistical Modelling of Extreme Values. Springer.

Colpan A. M. and Hikino T. (2010). "Foundations of Business Groups: Towards an Integrated Framework" in The Oxford Handbook of Business Groups, Colpan et al. (eds). Oxford University Press.

Dunning, J. H. (1977). "Trade, location of economic activity and the MNE: a search for an eclectic approach", in The international allocation of economic activity: proceedings of a Nobel Symposium held at Stockholm, Ohlin et al. (eds). The Macmillan Press Ltd.

Economist, The (2012) "Dead Money. Cash has been piling up on companies' balance sheets since before the crisis", in The Economist on 3rd November 2012.

Eurostat (2007). Recommendations Manual on the Production of Foreign Affiliates Statistics (FATS). European Commission.

Fan J. and Lang L. (2000). "The measurement of relatedness: An application to corporate diversification", Journal of Business 73, 629–660.

Fruin W. M. (2008). "Business Groups and Inter-firm Networks" in The Oxford Handbook of Business History, Jonathan and Jones (eds). Oxford University Press.

Garicano L. (2000). "Hierarchies and the Organization of Knowledge in Production", Journal of Political Economy, vol. 108(5), pp. 874-904.

Garicano, L. and Hubbard T. N. (2007)- "The Return to Knowledge Hierarchies.", NBER Working Paper N. 12815.

Garicano L. and Rossi-Hansberg E. (2006). "Organization and Inequality in a Knowledge Economy.", The Quarterly Journal of Economics, 121(4), pp. 1383-1435.

Granovetter M. (1995). "Coase revisited: Business Groups in the Modern Economy", Industrial and Corporate Change 4 (1), pp. 93-130.

Grossman S. and Hart O. (1986). "The costs and benefits of ownership: A theory of vertical and lateral Integration", Journal of Political Economy 94, pp. 691–719.

Grossman G. M. and Helpman E. (2002). "Integration versus Outsourcing in Industry Equilibrium", The Quarterly Journal of Economics 117(1), pp. 85-120.

Grossman G. M. and Helpman E. (2003). "Outsourcing versus FDI in Industry Equilibrium", Journal of the European Economic Association vol. 1(2-3), pp. 317-327.

Grossman G. M. and Helpman E. (2004). "Managerial incentives and the international organization of production," Journal of International Economics, vol. 63(2), pp. 237-262

Grossman G. M. and Helpman E. (2005). "Outsourcing in a Global Economy," Review of Economic Studies, vol. 72(1), pages 135-159.

Hart, O. and Moore J. (1990)."Property rights and the nature of the firm", Journal of Political Economy 98, pp. 1119–1158.

Helpman, E.(2006). "Trade, FDI, and the Organization of Firms", Journal of Economic Literature 44(3), pp. 589-630.

Hennart, J.-F. (1991). "Control in multinational firms : the role of price and hierarchy," Open Access publications from Tilburg University 12-174431, Tilburg University.

Hennart, J.-F. (1993). "Explaining the Swollen Middle: Why Most Transactions are a Mix of 'Market' and 'Hierarchy' ", Organization Science 4 (4) pp. 529-547.

Holmstrom, B. R., and Tirole J.(1989). "The Theory of the Firm", in Handbook of Industrial Organization, Schmalensee R and Willig R. (eds), vol. 1 (1). Elsevier.

Jones G.. and Colpan A. M. (2010). "Business Groups in Historical Perspective" in The Oxford Handbook of

Business Groups, Colpan et al. (eds.). Oxford University Press.

Joskow, Paul (2005). "Vertical integration", in Handbook of New Institutional Economics, Ménard C., and Shirley M. (eds). Springer.

Kali R. (1999). "Endogenous Business Networks", Journal of Law, Economics, and Organization 15(3), pp. 615-636.

Kali R. (2003). "Business Groups, the Financial Market and Modernization", Economics of Transition 11(4), pp. 671-696.

Khanna T. and Yafeh Y. (2007). "Business Groups in Emerging Markets: Paragons or Parasites?", Journal of Economic Literature vol. 45 (2): 331-372.

Kumar K., Rajan R. and Zingales L. (1999). "What determines firm size?", NBER Working Paper N. 7208, National Bureau of Economic Research.

Lafontaine F. and Slade M. (2007). "Vertical Integration and Firm Boundaries: The Evidence," Journal of Economic Literature 45(3), pp. 629-685

La Porta R., Lopez-de-Silanes F.and Schleifer A. (1999). "Corporate Ownership around the World", Journal of Finance 54(2), pp. 471-517.

Marin D. and Verdier T. (2009). "Power in the multinational corporation in industry equilibrium," Economic Theory, 38(3), pp. 437-464.

Melitz, M. J. (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", Econometrica, 71(6), pp. 1695-1725.

Nunn N. (2007). "Relationship Specificity, Incomplete Contracts and the Pattern of Trade", The Quarterly Journal of Economics , vol. 122(2), pp. 569-600

Nunn N. and D. Trefler (2008). "The Boundaries of the Multinational Firm: An Empirical Analysis", in E. Helpman, D. Marin, and T. Verdier (eds.), The Organization of Firms in a Global Economy, Harvard University Press, pp. 55-83.

OECD (2005). Guidelines for Multinational Enterprises. OECD.

Powell, W. W. (1990). "Neither Market nor Hierarchy: Network Forms of Organization", Research in Organizational Behavior 12, pp. 295-336.

Rajan R. G. and Zingales L. (1998). "Financial dependence and growth", American Economic Review, American Economic Association, 88, pp. 559-586

Rajan R. G. and Zingales L. (2001a). "The Firm As a Dedicated Hierarchy: A Theory Of The Origins And Growth Of Firms", The Quarterly Journal of Economics, vol. 116(3), pages 805-851.

Rajan R. G. and Zingales L. (2001b). "The Influence of the Financial Revolution on the Nature of Firms," American Economic Review vol. 91(2), pages 206-211.

Shannon, C. E.(1948) "A Mathematical Theory of Communication." The Bell System Technical Journal 27, 379-423 and 623-656

World Bank (2011a). Worldwide Governance Indicators. World Bank Group (available at: http://info.worldbank.org/governance/w

World Bank (2011b). Global Development Finance. World Bank Group (available at: http://data.worldbank.org/data-catalog/global-development-finance).

UNCTAD (2011). World Investment Report 2011. United Nations.

UNCTAD (2009). Training Manual on Statistics for FDI and the Operations of TNCs Vol. II. United Nations.

US BEA (2012). U.S. Multinational Companies. Operations of U.S. Parents and Their Foreign Affiliates in 2010. November 2012. (http://www.bea.gov/scb/pdf/2012/11%20November/1112MNCs.pdf)

Vitali S., Glattfelder J.B., Battiston S. (2011). "The Network of Global Corporate Control". PLoS ONE 6(10): e25995. doi:10.1371/journal.pone.0025995.

Whinston, Michael (2001). "Assessing the property rights and transaction-cost theories of firm scope", American Economic Review Papers and Proceedings 91, 184–188.

Williamson, Oliver (1971). "The Vertical Integration of Production: Market Failure Considerations", American Economic Review 61 (2), pp. 112-123.

Williamson, Oliver (1975). Markets and Hierarchies: Analysis and Antitrust Implications. Free Press, New York.

Williamson, Oliver (1985). The Economic Institutions of Capitalism. Free Press, New York.

Appendix A: Corporate control and Business Groups from the Ownership Database (BvD)

Our two main sources of data are both compiled by Bureau Van Djik (BvD), a consulting firm, and comprise the Ownership Database, from which we derive information on intra-group control linkages, and the Orbis database, from which we retrieve companies' balance sheet information.

The Ownership Database, in particular, includes information on over 30 million shareholder/subsidiary links for companies worldwide. Information on proprietary linkages is collected directly from single companies, from official bodies when in charge, or from some national and international providers. In Table A1 we include a list of the information providers, with the indication of the countries/areas they cover, as reported by the Ownership Database. In case of conflicting information among providers covering the same country/area, the Ownership Database is updated according to the latest available report.

Among the international providers, Bureau van Djik enlists also Dun & Bradstreet, a data source that has already been exploited in other academic works mentioned in this paper (Acemoglu, Johnson and Mitton, 2009; Alfaro et al., 2011; Alfaro and Charlton, 2009).

Table A1: Original sources of ownership linkages collected by Bureau Van Djik

CIBI Information, Inc. (Philippines), Creditreform (Bulgaria, Ukraine & Rep. of Macedonia) , Chamber of Commerce & Industry of Romania (Romania), CMIE (India), CFI Online (Ireland), Creditreform-Interinfo (Hungary), Infocredit Group Ltd, (Cyprus & Middle East), CreditInform (Norway), Creditreform Latvia (Latvia), Creditreform (Rep. of Macedonia), Informa Colombia SA (Colombia), Contact database, Credinform (Russia & Kazakstan), Creditreform Austria (Austria), Coface Slovenia (Slovenia), Dun & Bradstreet (USA, Canada, Latin America & Africa), DGIL Consult (Nigeria), MarketLine, (previously Datamonitor), PT. Dataindo Inti Swakarsa (Indonesia), DP Information Group (Singapore), Finar Enformasyon derecelendirme ve danismanlik hizmetleri A.S (Turkey), Suomen Asiakastieto (Finland), Factset, Worldbox (Switzerland), Honyvem (Italy), Creditreform Croatia (Croatia), Huaxia (China), Inforcredit Group (Cyprus), Informa del Peru (Peru), ICAP (Greece), Informa (Spain), InfoCredit (Poland), Ibisworld (Australia), Jordans (UK, Ireland), Patikimo Verslo Sistema (Lithuania), Krediidiinfo (Estonia), Købmandstandens Oplysningsbureau (Denmark), KIS (Korea), LexisNexis (Netherlands), Bureau van Dijk (Luxemburg), Creditreform Belgrade (Bosnia-Herzegovina, Serbia & Montenegro), Coface MOPE (Portugal), National Bank of Belgium (Belgium), Novcredit (Italy), Qatar Chamber of Commerce and Industry (Qatar), Annual return (UK), Coface SCRL (France), Creditinfo Schufa GmbH (Czech Republic, Slovakia, Iceland, Malta), SeeNews (Moldova, Albania, Georgia & Uzbekistan), Chinese source, Statistics Canada (Canada), China Credit Information Service Ltd (Taiwan), Taiwan Economic Journal (Taiwan), Teikoku Databank (Japan), Transunion (South Africa), UC (Sweden), Verband der Vereine Creditreform (Germany), Worldbox (New Zealand, Hong Kong, Switzerland, Monaco, Liechtenstein, Pakistan, Sri Lanka & Cuba)

The observation unit collected by the Ownership Database is the single link between a company and each of its shareholders, with additional information on the total (direct and indirect) equity participation when relevant. There are 7,707,728 companies with information on shareholding structures in the original database. An algorithm provided by Bureau van Dijk allows to identify in principle the ultimate owners of a single company.

However, since our purpose is to track the whole network of firms developed by each Business Group as defined in Section 2.1 and model it as a hierarchical graph (see Figure 1), we have in principle to depart from the complete shareholding structure of each company, in order to identify one ultimate parent company, its set of affiliates and their relative distance within the hierarchy. To that extent, we slightly modify the original BvD algorithm in two ways: we reconcile conflicting information that can come from controlling and controlled subjects and we differentiate between corporate and individual ultimate owners, recovering a total of 1,790, 062 firms which belong to Business Groups (270,474 parents and 1,519,588 affiliates) according to our definition.

Conflicting information deriving from controlling and controlled subjects can arise in presence of cross-participations. In accordance with international standards we apply a threshold criterion (>50.01%) for the definition of control on the basis of (direct and indirect) participation. The latter is the methodology currently used across international institutions (OECD 2005; UNCTAD, 2009; Eurostat, 2007), although it can lead to an overestimation of control in some bigger networks of affiliates.[35] That is, it is possible to end up with one affiliate controlled by more than one ultimate parent company even after adopting a majority threshold. To solve that problem we can rely on information officially provided by companies' consolidated financial accounts, when available. In particular, if we find that an affiliate is enlisted in more than one Business Group, we give priority to the ultimate parent company that enlists that affiliate in its consolidated accounts.

On the other hand, as the standard algorithm reports every property linkage between a company and each of its shareholders, it includes as members of potential business groups (as previously defined) also affiliates that are directly controlled by individual (non-corporate) shareholders, and that are not controlling subjects of any other company. While we have excluded these cases from our sample, we include in our analysis those corporate networks that involve at least one intermediate property linkage of a corporate nature.[36]

More in detail, our modified algorithm partitions all firms for which information on ownership is available preliminarly in two groups:

a) a set of independent companies, that have as controlling shareholder individuals or a family or no specific controlling entity, and that are not themselves controlling shareholders of any other company;

b) all the other companies for which information on property linkages is available; these companies are either owned by a corporate controlling (immediate) shareholder or are themselves independent, but act as controlling shareholders of other companies.

We exclude the set a) of independent firms from our sample, and use it as a control group for further empirical analysis (see below Table A3 for a description of this latter sample).

The algorithm then screen every firm belonging to group b) for the highest total (direct and indirect) participation in the equity of each company, as provided by the Ownership Database. Once it finds a corporate controlling entity A that sums up to more than 50.01% of control in a given company B, company B is classified as an affiliate, while the same algorithm checks the shareholding structure of company A. If the latter is in turn ultimately owned by another corporate entity C, the process is repeated until a controlling company that has no corporate controlling shareholder is found. The latter is considered as the ultimate parent company of affiliate companies A, B and C. In the case of quoted companies, we consider as ultimate parent the highest company in the path of proprietary linkages we can identify.

[35]See Vitali et al. (2011) for an assessment of existing methodologies to attribute global corporate control and their limits.

[36]If for example an individual X directly controls affiliates A and B, we do not consider the X-A-B network as a business group. Whereas, in the case of an individual X that indirectly controls affiliates C and D through a third company E, we consider the E-C-D network as a Business Group, in which company E is the ultimate (corporate) owner.

Having identified the set of affiliates and their parent, the algorithm then assigns a hierarchical level within each Business Group, counting from the parent how many steps of intermediate property are required for ultimate control. In case the same affiliate is encountered more than once in the same path (due to cross-participations), we consider it as located on the farthest level where we have first encountered it.

A limit of the Ownership Database of Bureau Van Dijk concerns the maximum number of control levels that can be obtained after considering cross-participations: the algorithm allows to reach a maximum of 10 levels for a maximum of 1,000 affiliates. However, in our data only 13 Business Groups (that is 0.005% of our sample) exceed such limits. For these groups we can still obtain balance sheet data for each affiliate and the headquarter, but we cannot retrieve the position of each affiliate on the control chain. Hence, these 13 groups will be excluded from the empirical analysis involving measurements of the hierarchical complexity.

As a result of our procedure to identify Business Groups, we can derive a control group that we employ for a preliminary comparison of affiliates to Business Groups and non-affiliated firms along different dimensions. From the above preliminary division in two groups of all firms for which we have data in the Ownership Database, the group a) consists of 6,084,115 firms controlled by individuals or families without control shares in any other company. Among them, however, for only 3,756,003 we can retrieve financial accounts that we use for the calculation of premia in Table 2.

In Table A2 we report some descriptive statistics of the control group collecting non-affiliated firms confronted with the set of affiliates belonging to Business Groups.

Table A2: A control group of non-affiliated firms, some descriptive statistics

	Non-affiliated firms					Affiliates of BGs				
Variables	**Mean**	**Median**	**St. dev.**	**Min**	**Max**	**Mean**	**Median**	**St. dev.**	**Min**	**Max**
Employment	13	4	51	1	14717	367	25	6225	1	1917456
Fixed assets	931	70	15016	1	4593066	109404	728	1812076	1	2.79E+08
Turnover (sales)	1464	279	4344	1	979128	119045	3493	1616124	1	3.57E+08

Appendix B:
Correspondence table between US input output table and NAICS 2002 classification of industries

Table B1: Correspondence table NAICS 2002 and US BEA Input Output tables

	Industry description	Input output codes	2002 NAICS codes
1	Farms	111CA	111, 112
2	Forestry, fishing, and related activities	113FF	113, 114, 115
3	Oil and gas extraction	211	211
4	Mining, except oil and gas	212	212
5	Support activities for mining	213	213
6	Utilities	22	22
7	Construction	23	23
8	Wood products	321	321
9	Nonmetallic mineral products	327	327
10	Primary metals	331	331
11	Fabricated metal products	332	332
12	Machinery	333	333
13	Computer and electronic products	334	334
14	Electrical equipment, appliances, and components	335	335
15	Motor vehicles, bodies and trailers, and parts	3361MV	3361, 3362, 3363
16	Other transportation equipment	3364OT	3364, 3365, 3366, 3369
17	Furniture and related products	337	337
18	Miscellaneous manufacturing	339	339
19	Food and beverage and tobacco products	311FT	311, 312
20	Textile mills and textile product mills	313TT	313, 314
21	Apparel and leather and allied products	315AL	315, 316
22	Paper products	322	322
23	Printing and related support activities	323	323
24	Petroleum and coal products	324	324
25	Chemical products	325	325
26	Plastics and rubber products	326	326
27	Wholesale trade	42	42
28	Retail trade	44RT	44, 45
29	Air transportation	481	481
30	Rail transportation	482	482
31	Water transportation	483	483
32	Truck transportation	484	484
33	Transit and ground passenger transportation	485	485
34	Pipeline transportation	486	486
35	Other transportation and support activities	487OS	487, 488, 492
36	Warehousing and storage	493	493
37	Publishing industries (includes software)	511	511, 516
38	Motion picture and sound recording industries	512	512
39	Broadcasting and telecommunications	513	515, 517
40	Information and data processing services	514	518, 519
41	Federal Reserve banks, credit intermediation, and related activities	521CI	521, 522
42	Securities, commodity contracts, and investments	523	523
43	Insurance carriers and related activities	524	524
44	Funds, trusts, and other financial vehicles	525	525
45	Real estate	531	531
46	Rental and leasing services and lessors of intangible assets	532RL	532, 533
47	Legal services	5411	5411
48	Computer systems design and related services	5415	5415
49	Miscellaneous professional, scientific, and technical services	5412OP	5412-5414, 5416-5419
50	Management of companies and enterprises	55	55
51	Administrative and support services	561	561
52	Waste management and remediation services	562	562
53	Educational services	61	61
54	Ambulatory health care services	621	621
55	Hospitals and nursing and residential care facilities	622HO	622, 623
56	Social assistance	624	624
57	Performing arts, spectator sports, museums, and related activities	711AS	711, 712
58	Amusements, gambling, and recreation industries	713	713
59	Accommodation	721	721
60	Food services and drinking places	722	722
61	Other services, except government	81	81

Appendix C: Sample properties of the Group Index of Complexity (GIC)

We had already argued in Section 5 how a simple counting of the number of affiliates can be a poor proxy for the complexity of the control chain developed by a Business Group, where the hierarchy design can involve different dimensions and groups with the same number of affiliates can eventually come to display very different proprietary structures. Hence, we borrowed from graph theory a notion of (probabilistic) entropy able to provide a synthetic measure for the similarity of hierarchical graphs once assuming that Business Groups' control chains can be conveniently represented with a vertex (here the headquarter) connected through edges (in our case control participations) to single nodes (represented by affiliates) for the purpose of a coordinated management of economic activities. Thereafter we argued that a change in the original node entropy was necessary to introduce an increasing marginal complexity when affiliates were added at a farther distance from the headquarter and we discussed how this alteration modified consequently the properties of the original measure.

In this Appendix we want to show with further details how our Group Index of Complexity (GIC) relates to both the number of affiliates (N) and the original node entropy (H(G)) borrowed from graph theory exploiting our dataset of Business Groups introduced in Section 3.

In Table C1 we report some descriptives that already show how both the node entropy in the second column and the GIC in the third colum reproduce long right-tail distributions similar to the more simple number of affiliates but with some differences. Skewness is much higher in the case of N, while H(G) and GIC start increasing rapidly only after the 83rd percentile of our sample, differently from the distribution of N which already begins a right tail from the 75th percentile.

Table C1: Comparison of distributions: number of affiliates, node entropy and GIC

Statistics	N	H(G)	GIC
Mean	5.62	0.18	0.35
standard deviation	32.62	0.43	1.02
Skewness	28.59	2.43	5.34
50th percentile	1	0	0
75th percentile	3	0	0
83th percentile	4	0.41	0.68
90th percentile	8	0.92	1.45
95th percentile	16	1	1.88
99th percentile	74	1.83	4.7
Maximum	1000*	3.27	19.07

*Maximum number of affiliates for which full proprietary linkages can be retrieved in our sample, excluding 13 BGs for which GIC and node entropy $H(G)$ cannot be calculated.

Since until the 57th percentile our sample is represented by Business Groups having only one affiliate, both the node entropy and the GIC end up with null figures until that point of the distribution.

However, given the logarithimic weight of formulas 3 and 4, also groups that have few affiliates but all positioned on a same proprietary level ($N = n_l$,whatever l) end ud with a H(G) and GIC that have final null figures until the 83rd percentile.

This last feature can be interpreted in terms of graph theory as a minimum complexity of the hierarchical graph when nodes are all adjacent on a same level. From an economic point of view it makes sense that Business Groups having affiliates all located at the same proprietary distance from the headquarter are more easily coordinated in the management of their activities. Moreover, if we assume that control runs univocally from the headquarter to each single affiliate, the cost of maintaining a control chain with only one affiliate and the cost of it where more affiliates are however located at the same level are virtually the same.

The previous is the reason why GIC (and node entropy) is not monotonic in the number of affiliates, since the GIC (and the node entropy) is additive in proprietary levels but not in number of affiliates. Groups with the same number of affiliates can arrange them in one or more levels and the cost of exerting control through the network is higher in the latter case.

In Table C2 we show however how sample distributions of GIC and N are significantly correlated when we report descriptives of the first by Business Groups'size classes

Table C2: Descriptives of GIC by size classes of Business Groups

N. affiliates	Frequency (%)	mean GIC	median	st dev	max
1-5	86.70	0.16	0.00	0.51	6.97
6-20	9.49	1.21	0.99	1.47	18.87
21-50	2.29	2.09	1.51	2.16	18.83
51-200	1.21	3.29	2.69	2.90	18.71
201-500	0.23	5.20	4.27	3.88	18.96
>500	0.09	5.86	5.25	3.89	19.07

The average GIC and its median in the second column is indeed increasing from small groups that have until five affiliates to bigger groups that report more than 500 affiliates. Standard deviations by size classes are also moderately increasing revealing that within each of them there is a certain (increasing) degree of heterogeneity of organizational complexity of affiliates along control chains, confirming that GIC is more able to catch this heterogeneity in the hierarchy design. Indeed, in the last column, where we report maximum GIC by size class, we do observe that even groups with only ten affiliates can show complex control chains comparable to bigger groups because, recalling entropy properties introduced in Section 5, the GIC grows rapidly when more affiliates are located on farther levels[37].

One would like to graphically compare N and GIC to observe how they behave along our sample distributions. Given their extremely long right tail, we resort to mean excess plots of Figure C1, where graphs are the results of the following so called excess mean function (see Beirlant et al., 1996):

$$e_F(u) = E\left(Y - u \mid Y > u\right) \tag{C1}$$

[37]More precisely, given N number of affiliates and L number of levels for each Business Group, the specific GIC reaches its maximum when $m \cdot \frac{n_m}{N} = n \cdot \frac{n_n}{N}$, for each $m, n \in L$, n_m and n_n affiliates on mth and nth level.

where $u \in (x_l, x_r)$ where $x_l = \inf\{x : F(x) > 0\}$ and $x_r = \sup\{x : F(x) < 1\}$.

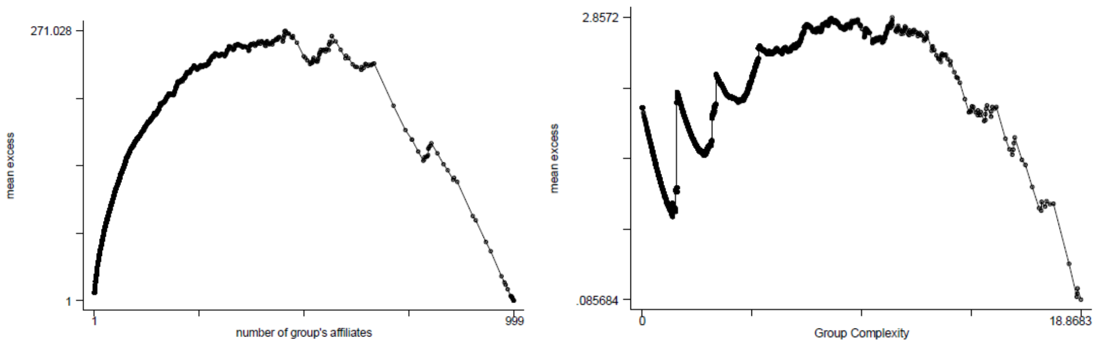
Shortly, the mean excess plot gives back a transformation of the distribution such that it is represented as the excess of the consecutive sample means of a random variable Y with respect to a sequence of threshold values u picked from the same domain of random variable Y such that the cumulative distribution $F(x)$ is defined on the interval $(0, 1)$. More analitically, we can describe the mean excess function as:

$$e_F(u) = \frac{1}{F(u)} \int_u^\infty F(y)dy \tag{C2}$$

with distributions of random variable Y on $[0, \infty)$.

In our specific case, mean excess plots in Figure C1 depart from the first threshold value u different from zero from our sample of Business Groups (1 in the case of N and 0.12 in the case of GIC). The values at which the mean excess functions are maximum are 271 for random variable N and 2.89 for random variable GIC.

Figure C1: A visual comparison of number of affiliates (N) and GIC distributions: mean excess functions



In both cases, we observe that after the maximum of the mean excess function, the distributions are steepily decreasing almost linearly, confirming that they have a long right-tail that is very much similar from a statistical point of view. However, on the left tail of the distribution, number of affiliates and GIC behave differently, where some points of discontinuity are present for the GIC excess function Those discontinuities reveal that at some lower measures of organizational complexity there is a high portion of Business Groups from our sample, a feature that was not observed when looking at the more simple group size in terms of number of affiliates.

Finally, in Figure C2 we report the quantile-quantile plot of GIC against N, in order to observe graphically how they behave with respect to each other by percentiles. The graph shows that for intermediate number of affiliates, GIC has a smoother distribution especially once approaching groups that have a higher number of affiliates, being able to give them a continuous measure that take into account their organization along the value chain. Business Groups with, say, about a thousand of affiliates can show a different GIC, however higher with respect to smaller groups if affiliates are located at increasing levels from the headquarter.

Figure C2: A quantile-quantile plot of number of affiliates (N) vs GIC