

Firm Performance and the Optimal Organization of Business Groups

Carlo Altomonte
(Bocconi University)

Armando Rungi
(IMT Lucca & FEEM)

May 21, 2015

Abstract

We exploit a unique dataset of 270,474 headquarters controlling more than 1,500,000 (domestic and foreign) affiliates in all countries worldwide in order to explore the relation between the organization of a business group and its performance. We study two different organizational choices of BGs, that is the degree of their vertical integration, and the hierarchical complexity of their structure of control, focusing on their interplay. We then relate these organizational choices to the productivity of the affiliates operating in the group. Controlling for the endogenous emergence of business groups as stemming from the institutional characteristics of the host country in which they operate, we find that vertical integration and organizational complexity tend to be positively correlated at the group level, but have a different impact on the specialisation of each affiliate within the group. We also corroborate the view that vertical integration and productivity tend to be correlated, but we find the relation not robust to the inclusion of a group's hierarchical complexity, as in our data the latter relation dominates the former. The positive and robust relation between hierarchical complexity and productivity is however non-linear: above a certain threshold of complexity the relationship becomes negative. Results are in line with the delegation view of business groups and the idea of knowledge-based hierarchies, in which intangible assets are a complementary input in the production processes.

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

Keywords: business groups, financial development, vertical integration, corporate ownership, organization of production, productivity.

**Financial support was provided by the Volkswagen Foundation, Programme "Europe's Global Linkages". We thank Sónia Araújo, Richard Baldwin, Giorgio Barba Navaretti, Paola Conconi, Wilhelm Kohler, Gianmarco Ottaviano, the team of the OECD Statistics Directorate and seminars audiences for useful comments at various stages of this work.*

1 Introduction

The emergence of Business Groups (BGs) is traditionally considered as an organizational response to circumvent market imperfections (Leff, 1978; Khanna and Palepu, 1997): 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 that mimics a market structure.¹ Hence, the emergence of such an organizational form has been traditionally correlated in the literature to differences in institutional quality across countries, in particular in terms of financial development (Belenzon et al., 2013a). What has been relatively less explored in the literature, however, is whether differences in institutional quality across countries, on top of their effect on the emergence of business groups, also shape their organization.² In turn, it is also unclear to what extent differences in the organizational choices of business groups relate with the performance of their affiliates. In this paper we aim at exploring systematically these linkages.

We capitalize on a unique dataset that we have built, able to map 270,374 headquarters controlling 1,519,588 affiliates worldwide in 2010, across all industries. The dataset is retrieved from ORBIS, a global dataset containing detailed balance sheet information for some 100 million companies worldwide, together with information on over 30 million shareholder/subsidiary links that we have been able to organize in a hierarchical control structure for each firm.³

Based on these data, we build two novel metrics aimed at assessing the organizational choices of BGs. We first collect information on the set of industries in which any affiliate of a group is engaged and (using an Input-Output matrix) combine this information with the input coefficient requirements that link those industries, in order to retrieve an index of vertical integration of the business group: the index measures, for every unit of output of the group, the percentage of inputs that can be potentially sourced from the affiliates of the same group.⁴ Endowed with the entire (direct and indirect) control structure of each group, we then combine information on the number of hierarchical levels of control and the density of affiliates operating at each level, retrieving an index of hierarchical complexity of a business group: the index is formalized as a variation of the node entropy of a hierarchical graph (Emmert-Streib and Dehmer, 2007), is continuous and additive in the number of levels / affiliates, and goes from zero (the simplest group, in which one parent controls one affiliate) to infinity, on a log scale. We then relate the metrics of BG's organizational choices to the productivity of the affiliates operating in the same group, which we can measure thanks to the availability of balance sheet data

¹Across 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.

²Belenzon et al. (2013a) show that, even within their relatively homogeneous sample of European countries, there is substantial variation in the organizational form of business groups.

³Belenzon et al. (2013a) use the European version of ORBIS, known as Amadeus, in their analysis of European Business Groups (with 2007 data). On top of European data, the ORBIS 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.

⁴In this, we improve on Acemoglu et al. (2009), who have built a similar index but have treated each firm in their sample as independent, thus neglecting the possibility that the degree of vertical integration can be a function of the coordinated management decision within a Business Group, where the decision to "make or buy" can be differentiated between headquarters and affiliates or across the same affiliates.

for (most of) the firms in our dataset.⁵

To better fix the ideas on how we measure differences in the organizational choices of groups, and how the latter possibly relate to productivity, consider the case of two business groups present in our data: 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, these two groups have a similar size, as they control directly or indirectly around 650 affiliates each (2010 data).⁶ The geographical distribution of affiliates (and hence the institutional quality of the countries in which they are located) differ for the two groups: 65% of affiliates of Mitsubishi are concentrated in Japan, with relatively large numbers also in the US and UK (some 10% of affiliates) and then a presence of the group in 29 other countries, mostly Asian; on the contrary, GM has a 44% share of domestic (US) affiliates, some 20% of affiliates in the UK, but then a more diffused international presence in 52 other countries. When looking at industrial activities beyond motor vehicles, Mitsubishi is involved 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. In our metrics, the latter results in a higher degree of vertical integration for Mitsubishi than GM. Still, Mitsubishi is characterized by a much flatter control structure (3 levels of hierarchy within the group), while GM is characterized by a deeper (up to 8 levels) hierarchy of cross-participations, with affiliates present with similar shares across different levels of control. In our metrics, the latter implies that GM has a more complex hierarchical structure than Mitsubishi. Yet, when looking at performance, the average productivity of affiliates belonging to the hierarchically more complex (but less vertically integrated) GM group is significantly higher than within Mitsubishi.

The latter evidence of a complex interplay between vertical integration choices, the complex organization of the hierarchy and performance (productivity) is indeed systematic and statistically significant across groups in our data. More specifically, when looking at a BG's organization, we find that vertical integration and organizational complexity tend to be positively correlated at the group level, but have a different impact on the specialisation of each affiliate within the group: conditional on the level of institutional quality of a host country and the overall organization of the group, affiliates tend to trade-off their degree of vertical integration with their position in the group hierarchy.

When looking at the correlation between a BG's organizational choice and performance, we cor-

⁵Not all firms in our dataset report a complete set of financial data and, moreover, country-level data for some institutional variables we use as controls are not available for every country. Hence, in our estimations we rely on a restricted sample of data (encompassing around 75% of the original firms) in which both balance sheet and country-level information are available. In the data section we discuss the validation of our general dataset as well as its comparison with the restricted dataset used in our econometric estimations.

⁶Alfaro 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*). Moreover some major events have occurred to GM since 1999. In 2005 the group conclusively sold its participations in electronics 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.

roborate the view that vertical integration and productivity tend to be correlated; we find however the relation not robust to the inclusion of a group’s hierarchical complexity, as in our data the latter dominates the former in terms of productivity. In general, we find a positive relation between hierarchical complexity and productivity; on top of our metric, the relation is also robust to the use of rougher proxies of complexity, such as the total distance (in terms of hierarchical levels) of each affiliate from the headquarter, or the total number of firms operating in the group. When using our proxy, however, we are able to discover more nuanced effects: first, the relation appears to be non-linear above a certain (albeit very high) threshold of complexity, i.e. the productivity of affiliates appears to be lower for groups having more than 550 affiliates and/or more than 5 levels of control. Such an evidence of marginally decreasing returns from increasing complexity is relevant, as it puts a natural limit to the growth in complexity of global Business Groups. Second, once controlling for the overall complexity of the business group, we find that the distance of an affiliate from the headquarter is (at the margin) negative related to its productivity.

We interpret all these findings as a refinement of the delegation view of corporate group structure (Belenzon et al., 2012): on the one hand, sequential delegation of authority allows for the exploitation of economies of scale for knowledge inputs, and hence affiliates belonging to bigger (more complex) networks are relatively more productive.⁷ At the same time, the non-linearity in the relationship shows that additional layers of control that foster the acquisition and communication of knowledge throughout the hierarchy are also associated to the emergence of endogenous communication costs (Caliendo and Rossi-Hansberg, 2012).

The paper is organized as follows. In Section 2 we introduce our dataset, providing at the same time some stylized facts. In Section 3 we construct the metrics of vertical integration and hierarchical complexity and describe their properties. Section 4 relates the same metrics 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

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

⁷A recent economic literature in particular models firms as knowledge-based hierarchies (Garicano, 2000; Garicano and Hubbard, 2007; Garicano and Rossi-Hansberg 2004, 2006, 2012): coordinated management decisions are taken on the basis of the available knowledge, considered as an intangible and costly input which is complementary to physical inputs in production processes. An organizational structure is hence endogenous and dependent on the costs of acquiring and communicating knowledge among agents involved with different tasks within the firm hierarchy. The positive correlation between hierarchical complexity and productivity is also consistent with the findings by Bloom, Sadun and Van Reenen (2012), according to which more delegation of authority implies a higher firm-level performance thanks to a better reallocation of resources.

⁸Business 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, 1993).

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 Wolfenzon, 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).

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. We thus follow Belenzon et al. (2013) and rely on an ownership-based definition of 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 impending hierarchies) and from multidivisional firms (which are organized through internal hierarchies of branches, but without autonomous legal status).⁹

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.

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 parent and one affiliate, or very complex groups with hundreds of domestic and/or foreign affiliates linked by hierarchical control. Hence, for the sake of generality, we rule out any *ad hoc* definition in terms of minimum number of affiliates or industries, as found in some part of the literature (e.g. Colpan and Hikino, 2010).

2.2 Data

The main difficulty in identifying BGs in the data is related to the notion of control exerted by a parent on affiliates via ownership. We opt here for a definition of control as established in international accounting standards (OECD 2005; UNCTAD, 2009; Eurostat, 2007), where control is assumed if

⁹The 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.

(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.¹⁰

Such a notion of control is not exhaustive, as it leaves outside the boundaries of BGs affiliates *de facto* controlled through minority ownership (<50%), 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 a definition of control allows for a straightforward comparison with official statistics, as the majority of voting rights is the criterion commonly used in international standards on foreign affiliates (Eurostat or OECD FATS) and for international tax purposes (IAS, IFRS).¹¹

Figure 1: Business Groups as hierarchical graphs

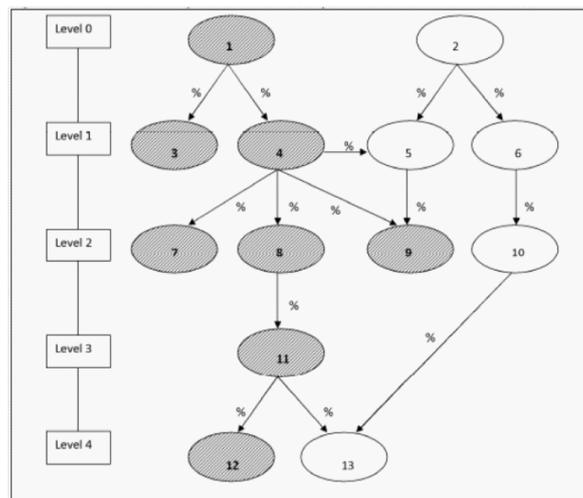


Figure 1 represents the organization of a typical Business Group as it can be derived by the application of the majority ownership notion of control that we use in this paper. Such a representation

¹⁰Control 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.

¹¹A 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, Glattfelder and Battistoni (2011) for a reference on this issue.

corresponds to a mathematical object known as a hierarchical graph.¹² The upper shaded node (1) represents the headquarters (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.¹³ The white nodes are instead firms possibly participated by the considered Business Group, but excluded from its boundaries on the basis of the majority ownership threshold.

Specifically, two different sorts of data have been combined to retrieve our dataset on 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.¹⁴ Both proprietary linkages and financial data refer to the last available information available in year 2010. Appendix A provides a detailed description of our data sources and of the methodology employed 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 and the number of co-affiliates operating at the same level of control. This information is then used to build our measure of organizational complexity. For each headquarter and each affiliate along the control chain we have also recovered industry affiliations at the 6-digit NAICS rev. 2002 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 our proxies of 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 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 balance sheet 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 of the data described in this section also hold for the restricted sample of Business Groups.

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 for all countries of the world. One partial exception is the World Investment Report of UNCTAD, which compiles yearly a list of the biggest corporations currently operating in the world: these firms are all present in our dataset, with their affiliates duly recorded. UNCTAD (2011) also reports detailed

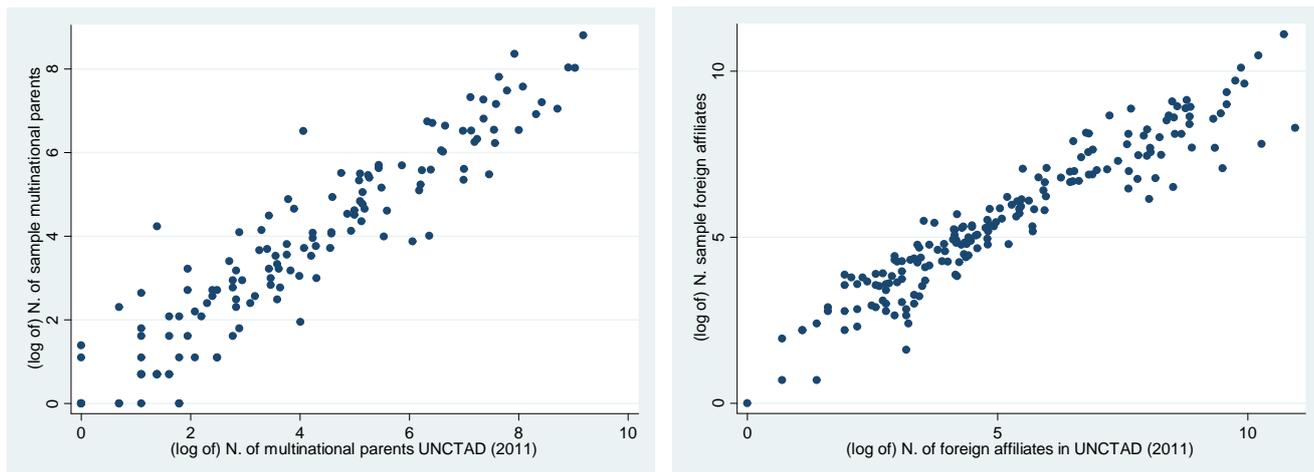
¹²Technically, a hierarchical graph is a particular variation of a flat graph to which at least one parent node is added so as to assign functions to the other nodes (Palacz, 2004). Hierarchical graphs in turn are a generalization of a tree graph, in which several arms depart from one vertex as in a 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 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 with its command chain.

¹³In 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.

¹⁴The data sources are similar to Belenzon et al. (2013), with the difference of an extended worldwide coverage, for the year 2010. Past studies in the literature (e.g. Alfaro et al, 2009) have also exploited 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.

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 and matched against the corresponding figures provided by UNCTAD (2011): correlations are .94 and .93, respectively.¹⁵

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



2.3 Stylized facts on 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 these headquarters 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 total number of foreign affiliates located in a given country / area, as resulting from inward FDI.

Two thirds of our BGs' parent companies are headquartered in OECD economies, then controlling about 76% of affiliates worldwide. The ratio of foreign to domestic affiliates is smaller for groups originating from developing countries (around .26, i.e. one foreign affiliate each four domestic ones on average), since these countries have a relatively larger proportion of firms organized as domestic business groups (Khanna and Yafeh, 2007). The ratio is instead highest for the US (.85), where in particular 32% of Business Groups are only domestic, 24% are only cross-border with all affiliates abroad, while 44% tend to operate both domestically and abroad.

¹⁵The 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 1: Geographic coverage of Business Groups (main countries/areas) by headquarters and affiliates

Economy	N. of parents (Business Groups)		Domestic affiliates (A)	Affiliates abroad (B)	Foreign affiliates located in economy
	N. of affiliates (A + B)				
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
<i>of which:</i>					
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

Only selected countries/areas are reported. Totals refer to all countries present in the complete sample.

Looking at individual balance sheet data, we find that very simple Business Group structures (with one headquarter and one affiliate, located either domestically or abroad) represent 57% of our groups, but account for only 1% of the total value added in our data. On the contrary, around 2,000 of the largest BGs (headquarters with more than 100 affiliates, both domestic and cross-border) constitute less than 1% of groups in our sample but account for 72% of the total value-added measured in the data. In the US, these large ‘mixed groups’ operating both domestically and abroad report a value added equal to 83% of the total US value-added recorded in the sample.

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 characteristic 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 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

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*** (.008)	.02*** (.008)	.05*** (.008)

Binary regressions with country-per-industry fixed effects; **, *** stand for significance respectively at 5% and 1%; (1)

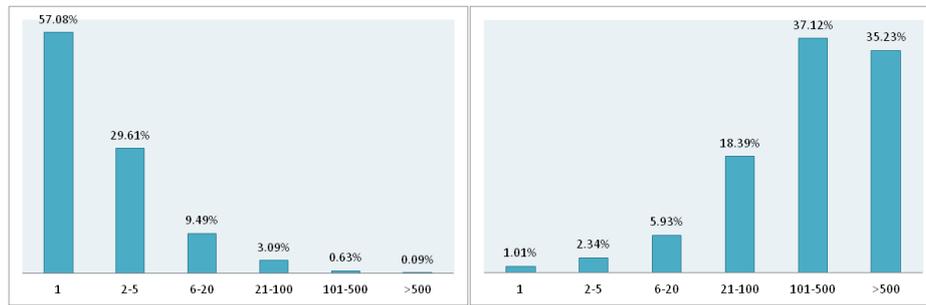
Capital-intensity and size added as a further control for a one-factor measure of productivity. See Appendix A for details on the control group of non-affiliated firms.

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.¹⁶ 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 sections 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.

¹⁶This 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)

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
USA	21	3	9	92	354	2,707
Rest of the world	5	1	3	15	60	1,672
<i>of which:</i>						
Japan	8	1	4	31	119	2,534
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

3 Organizational Metrics for Business Groups

3.1 Group vs. Affiliate Vertical Integration

Acemoglu, Johnson and Mitton (2009, henceforth AJM) have explored 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.¹⁷

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.¹⁸

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

To take into account the latter dimension, we have thus 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 activities 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 activities 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 i th coefficient requirement to produce the j th output.

¹⁷They 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.

¹⁸For 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).

Figure 4: A group-specific input-output table

		Outputs - headquarter's activities					Intermediate demand by input
		j_1	j_2	j_3	...	j_{δ_1}	
Inputs - affiliates' activities $i \in N_A$	i_1	$VI_{i_1j_1}$	$VI_{i_1j_2}$	$VI_{i_1j_3}$...	$VI_{i_1j_{\delta_1}}$	$\sum VI$ from input i_1
	i_2	$VI_{i_2j_1}$	$VI_{i_2j_2}$	$VI_{i_2j_3}$...	$VI_{i_2j_{\delta_1}}$	$\sum VI$ from input i_2
	i_3	$VI_{i_3j_1}$	$VI_{i_3j_2}$	$VI_{i_3j_3}$...	$VI_{i_3j_{\delta_1}}$	$\sum VI$ from input i_3

	i_{δ_1}	$VI_{i_{\delta_1}j_1}$	$VI_{i_{\delta_1}j_2}$	$VI_{i_{\delta_1}j_3}$...	$VI_{i_{\delta_1}j_{\delta_1}}$	$\sum VI$ from input i_{δ_1}
	Intermediate usage by output	$\sum VI$ for output j_1	$\sum VI$ for output j_2	$\sum VI$ for output j_3	...	$\sum VI$ for output j_{δ_1}	

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 tables¹⁹ 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 rev. 2002 classification. In Appendix B we report the official 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 integration 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:

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

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

¹⁹As in AJM(09), the use of the US input-output table for all countries is justified by the assumption that there is a correlation in the input use patterns across countries. More in general, at the basis of the use of a common input-output table there are the assumptions of a common technology frontier and either of a Leontief production function or of factor price equalization.

²⁰As 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.

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 discussed in the Introduction of the paper: the American group is relatively specialized, 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 active in more diversified sourcing industries than 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 belonging to the same group, thus generating potentially biased results. This is an important feature of Business Groups’ 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 headquarters, 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:

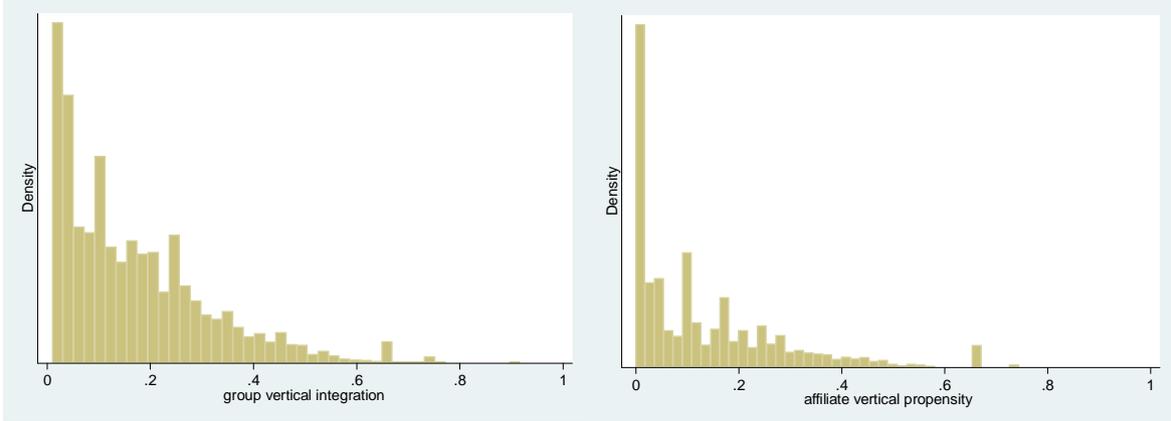
$$v_a = \sum_{\substack{i \in N_a \\ j \in N_H}} \frac{1}{|N_a|} VI_{ij} \quad (2)$$

where the input coefficient requirements (VI_{ij}) are taken for any i th among single affiliate activities ($N_a \subset N_A$) that can lend to any j th 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 indexes 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 indices.

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. Similarly to Alfaro et al. (2011), both distributions

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



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

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

of our vertical integration indexes show long right-tails. In our case about 1.3% of Business Groups can potentially source internally more than 50% of the value of their output, while only about 0.8% of groups have affiliates that, individually taken, have vertical integration indexes in excess of 0.5.

3.2 Hierarchical complexity

A particularly convenient property of representing Business Groups as hierarchical graphs, as in Figure 1, is that it is possible to provide a synthetic measure of their organization through some hierarchical form of entropy. We can thus proxy the process of coordinated management that occurs within the hierarchy of firms in a BG by exploiting the information on the command chain that links single affiliates to the ultimate headquarter.

Borrowing from graph theory, the entropy of a hierarchical graph G characterized by a total of L levels of hierarchies can be constructed 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

$$H(G) = - \sum_l p_l \log(p_l) \quad (3)$$

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

²¹Defining $p_l = \frac{n_l}{N}$ 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 $n_l > 1$.

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 from 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:

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

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).

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 when affiliates are added at the same hierarchical level is associated to the idea that some economies of scale intervene when firms expand their network of affiliates horizontally, while coordination (and communication) costs can become more and more important once the network enlarges and deepens by 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

²²This 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.

scale are assumed, but at the same time it also increases the cost of communication along the hierarchy.

Accordingly, in our case the 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’ of knowledge. As a result, the hierarchical complexity of an object such as a Business Group cannot simply be proxied by its total number of affiliates N or by its number of hierarchical levels, with the index of complexity being not strictly monotonous in N . In Appendix C we provide evidence of a comparison between complexity proxied through a BG’s number of affiliates and our index of complexity.

3.3 Vertical Integration and Hierarchical Complexity across Countries and Industries

In Table 4 we report sample averages of both the Group Index of Complexity (GIC) and the group-level vertical integration (VPI), for some selected industries and geographical areas. The industry is identified as the core sector where the majority of value added is created within the Business Group, even though many larger BGs can be involved in more than one line of business. The country is instead the home country where the headquarter is located, even though the group can have some affiliates abroad.

The third and fourth columns of Table 4 show that, while group-level vertical integration is constantly lower for OECD economies with respect to non-OECD economies for each reported industry, the opposite is true for hierarchical complexity. Groups originated in the US are the ones showing higher figures for hierarchical complexity in most industries, while Japanese groups display instead lower delegation of control (they are hierarchically less complex). The figures for developing economies show instead a higher variation across industries.

As expected, the less integrated among the reported industries is the category of business services which rely less on physical inputs, while the most integrated groups can be found in the chemical industry. The automotive industry, from which we derived the case studies of General Motors and Mitsubishi sketched in the introduction, appears to be relatively less integrated than expected thanks to the presence of some very specialized small groups active in the provision of parts and components. Indeed, looking at the automotive industry in US and Japan, the preliminary evidence of the case studies is confirmed on industry aggregates, since the hierarchical complexity of the US car industry is higher than Japan. However, figures for vertical integration suggest that Japanese automotive groups exchange intermediates internally on average four times more than the US ones do.

Overall, cross-country variation seems to dominate cross-industry variation, especially when looking at figures of hierarchical complexities. Based on this evidence, in the next section we explore the relationship between group boundaries and country-level institutional determinants, controlling for the residual sectorial heterogeneity via fixed effects.

Table 4: Group vertical propensity and organizational complexity (averages) by selected industries and countries

Industry	Index	Countries/areas									
		OECD	non-OECD	European Union	USA	Japan	China	Africa	South America	ASEAN	All countries
Mining	group integration (v_g)	0.034	0.140	0.031	0.058	0.061	0.013	0.051	0.112	0.021	0.073
	GIC	0.626	0.356	0.124	0.764	0.339	0.370	0.597	0.924	0.922	0.530
Food	group integration (v_g)	0.071	0.151	0.065	0.105	0.138	0.067	0.078	0.074	0.083	0.114
	GIC	0.650	0.216	0.561	1.537	0.299	0.272	0.760	0.210	0.818	0.425
Textiles and clothing	group integration (v_g)	0.061	0.101	0.062	0.077	0.052	0.072	0.064	0.046	0.080	0.079
	GIC	0.503	0.170	0.426	1.639	0.257	0.134	1.160	0.366	0.836	0.349
Chemical products	group integration (v_g)	0.150	0.204	0.153	0.175	0.152	0.108	0.088	0.150	0.128	0.172
	GIC	0.789	0.278	0.538	1.579	0.455	0.096	0.301	0.293	0.736	0.588
Automotive	group integration (v_g)	0.067	0.069	0.044	0.137	0.186	0.033	0.067	0.120	0.081	0.068
	GIC	0.944	0.424	0.775	2.169	0.501	0.243	2.340	0.113	1.484	0.746
Electronic products	group integration (v_g)	0.079	0.114	0.059	0.115	0.102	0.104	0.129	0.044	0.068	0.094
	GIC	0.736	0.276	0.655	1.032	0.651	0.182	0.950	0.103	0.733	0.537
Business services	group integration (v_g)	0.013	0.036	0.012	0.048	0.071	0.080	0.100	0.101	0.042	0.022
	GIC	0.531	0.205	0.530	0.833	0.261	0.482	0.348	0.073	0.854	0.426
All sectors	group integration (v_g)	0.025	0.126	0.025	0.060	0.023	0.052	0.075	0.132	0.047	0.062
	GIC	0.418	0.233	0.410	0.989	0.114	0.311	0.601	0.308	0.808	0.354

Sample averages for group integration (v_g) and Group Index of Complexity (GIC). Industries are identified as the core activity where most value added is created. Countries as the origin of the parent company.

4 Empirical results

4.1 Group organization and institutional quality

We begin our analysis by applying the empirical strategy developed by Acemoglu, Johnson and Mitton (2009) to our group-specific measure of vertical integration in order to check to what extent institutional characteristics of the host countries in which the BGs operate are correlated with their organizational choices, also controlling for hierarchical complexity. We then replicate the analysis looking at how hierarchical complexity is correlated to institutional characteristics, also controlling for vertical integration. We finally look at how both institutional characteristics, vertical integration at the group level and hierarchical complexity interplay with vertical integration choices at the affiliate level.

In particular we assume that 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.²³ 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.²⁴ We thus test for the drivers of Business Groups' boundaries

²³At 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.

²⁴Although the latter is obviously a simplification 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

in nested steps: first we consider the drivers of group-level vertical integration; then we test for vertical integration at the affiliate level, given the choice of vertical integration at the group-level; further, we control for the level of hierarchical complexity of the group to which the affiliate belongs.

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 :

$$v_{gkc} = \beta_0 + \beta_1 X_{c_g} + \beta_2 Z_{c_g} + \beta_3 X_{c_g} Z_{c_g} + \beta_4 GIC_g + \beta_5 mne_g + \beta_6 \ln emp_g + \beta_7 \ln gdp_{c_g} + \gamma_k + \varepsilon_{gkc_g} \quad (5)$$

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 claim²⁵ and the country-level ratio of private credit provided by all financing institutions to GDP.²⁶

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 headquarters' balance sheet consolidated data, if available, or calculated summing up the employees of the headquarters and affiliates. 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 whether or not 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 (gdp_{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 one of the activities which we have used as outputs 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 5.

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:

had already been manufacturing.

²⁵The 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.

²⁶Private 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).

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

Dependent variable :	I	II	III	IV
Group integration				
contract enforcement	-0.139***		-0.114***	-0.116***
	(.037)		(.037)	(.037)
financial development		-0.085***	-0.070**	-0.071**
		(.035)	(.027)	(.028)
contract enforcement*financial development			.023	.020
			(.024)	(.024)
group index of complexity				.073***
				(.024)
multinational				-0.056
				(.035)
(log of) group employment	-0.003	.001	.003	.003
	(.006)	(.006)	(.004)	(.003)
(log of) GDP per capita	-0.234***	-0.229***	-0.188***	-0.185***
	(.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, errors clustered by country.

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 institutions 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.²⁷

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 5 confirms, which further strenghtens the idea that the home country institutional environment is a powerful driver of the organization of a Business Group.

We then replicate the analysis

[Paragraph + Table with the results of GIC as dep var vs. Institutions to be inserted here, with specification as in the previous Table, ie progressively insterting variable with

²⁷Indeed, in the robustness and sensitivity checks we present in Table 8, 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. Exactly as in their case, we also get correctly signed but poorly significant coefficients.

Table 6: Group index of complexity, vertical integration and institutional constraints

Dependent variable :	I	II	III	IV
Group index of complexity (GIC)				
contract enforcement	.009***		.011***	-.001
	(.001)		(.001)	(.001)
financial development		.006***	.008***	.013***
		(.001)	(.001)	(.001)
contract enforcement*financial development			-.001	.016***
			(.001)	(.001)
group vertical integration				.014***
				(.001)
multinational				.272***
				(.002)
(log of) group employment	.035***	.035***	.035***	.025***
	(.002)	(.006)	(.004)	(.003)
(log of) GDP per capita	.030***	.019***	.022***	.001
	(.001)	(.001)	(.001)	(.001)
Constant	-1.079***	-.970***	-1.001***	-.808***
	(.011)	(.014)	(.014)	(.014)
3-digit industry fixed effects	Yes	Yes	Yes	Yes
Errors clustered by country	Yes	Yes	Yes	Yes
Observations (N. of Business Groups)	221,701	221,701	221,701	221,701
Industries	88	88	88	88
Countries	129	129	129	129
Adjusted R squared	.100	.100	.101	.176

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

Given the ability of Business Groups to design vertical integration also across affiliates, we finally 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 g th group ($v_{a(g)kc_a}$), defined in Equation (2) 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

Table 7: Affiliate-level integration, group complexity and institutional constraints

Dependent variable :	I	II	III	IV	V	VI
Affiliate integration						
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 index of complexity						-.012
						(.014)
group index of complexity*contract enforcement						.019
						(.015)
group index of 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, errors clustered by country.

Business Group, in which case we will be dealing with a foreign affiliate.²⁸ Hence, the set of proxies of institutional frictions (X_{ca} , Z_{ca}), their interaction and the (log) of GDP per capita (gdp_{ca}) now all 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

²⁸In Table 8 we will specifically test for the robustness of our results when dealing with foreign affiliates only.

Table 7.

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 7, we introduce as a further control our metric of hierarchical complexity and its interaction with institutions. Results point at a potential trade-off between integration and complexity in the organizational design of a Business Group's boundaries: controlling for the level of financial development in the host country, a group's vertical integration and a group's hierarchical complexity have a different impact (one negative, the other positive) on the degree of vertical integration of each affiliate. In other words, controlling for institutions, a given level of integration achieved by affiliates stems from an alternative combination of vertical integration and hierarchy in the group. These results are in line with the case study of GM vs. Mitsubishi, in which different degrees of affiliates' integration can be associated to similar levels of group integration only since hierarchical complexity varies across the two groups (with GM's affiliates being placed in a relatively more complex hierarchical structure).²⁹

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 levels of vertical integration (e.g. considering all firms as independent in a sample) might thus lead to an omitted variable bias.

In Table 8 we present some tests of sensitivity of our results on the BG affiliates' organizational choices. 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

²⁹Clearly both group-level 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.

as well as its position in the hierarchy of the BG, 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.

Table 8: Sensitivity of results: affiliate-level vertical integration

Dependent variable :	AJM_09	> 20 employees	> 1 affiliate	domestic affiliates	foreign affiliates
Affiliate integration					
contract enforcement	-.042* (.024)	-.104** (.041)	-.094** (.042)	-.106** (.041)	-.054* (.026)
financial development	-.012 (.025)	-.151*** (.037)	-.149*** (.037)	-.152*** (.038)	-.101** (.046)
contract enforcement*financial development	.003 (.017)	-.003 (.028)	.003 (.030)	.024 (.031)	-.022 (.032)
group integration		.068*** (.015)	.061*** (.013)	.075*** (.025)	.030*** (.008)
group index of complexity		-.012 (.014)	-.016 (.014)	-.008 (.014)	-.014* (.006)
group integration*contract enforcement		-.020** (.008)	-.015* (.008)	-.013 (.012)	-.006* (.002)
group integration*financial development		-.064*** (.006)	-.064*** (.006)	-.081*** (.011)	-.028*** (.006)
group index of complexity*contract enforcement		.018 (.015)	.015 (.015)	-.005 (.015)	.004 (.006)
group index of complexity*financial development		.033*** (.011)	.038*** (.011)	.015** (.007)	.040* (.018)
(log of) group employment		-.001 (.004)	-.002 (.003)	.005 (.004)	-.001 (.001)
(log of) affiliate employment	-.003 (.004)	-.012 (.008)	-.010 (.008)	-.012 (.009)	-.003 (.007)
(log of) GDP per capita	-.086** (.037)	-.088 (.068)	-.096 (.069)	-.130** (.064)	-.013 (.070)
Constant	.872** (.376)	.908 (.715)	.979 (.725)	1.321* (.672)	.046 (.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, errors clustered by country.

In the second column of Table 8 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 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 8. 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, with respect to the average affiliate in our sample, affiliates of multinational corporations might be more influenced by organizational strategies developed in their country of origin, and hence relatively less influenced in their organizational design by the host country characteristics measured in this exercise.

4.2 Group boundaries and performance

Another crucial dimension in the analysis of Business Groups we look at in this paper is the relationship between the group’s organizational choices and the associated productivity of their affiliates. 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 reflects a cherry-picking effect, with already more efficient plants integrated *ex post* into more vertical structures.³⁰

Consistently with these findings, in Table 2 we have already 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, as previously discussed, vertical integration for Business Groups is a multi-dimensional concept (at the headquarters vs. affiliates’ level, and across affiliates distributed along more or less complex hierarchies).

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 g th group, operating in core industry k and located in country c_a .

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

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 another (simpler) proxy of hierarchical complexity. 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 correlation between organizational design and affiliates’ performance. Errors are clustered at the headquarter level, to account for within-group correlation. Nested results are presented in Table 9.

Looking at results, when we do not control for country fixed effects, as in the first column of Table 9, we obtain a negative correlation between vertical integration and productivity, both at the

³⁰For other works showing a positive correlation between vertical integration and productivity, see for example Maksimovic et al. (2002) or Schoar (2002), both investigating conglomerate firms. In the first a selection effect is claimed to be responsible for the productivity differences between integrated and non-integrated firms. In the second work diversification is a driver for such differences.

group- and at the affiliate-level. This is due to the global nature of our data: as also shown by Table 4, 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 9), we find the well-established positive correlation between our 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 9: 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 headquarters.

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).

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 only hierarchical complexity remains 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. Differently from Atalay et al. (2012), in our strategy we distinguish between (affiliate or group) size, measured by (affiliate or group) employment in the above specifications, and hierarchical complexity, being able to show that also this latter organizational dimension (on top of proxies of group size) correlates with affiliates' productivity.³¹

The robust finding that the positive correlation between hierarchical complexity and productivity dominates the one between vertical integration and productivity is in line with the delegation view of corporate group structure (Belenzon et al., 2012): sequential delegation of authority allows for the exploitation of economies of scale for knowledge inputs, and hence affiliates belonging to bigger (more complex) networks are relatively more productive. More specifically, the latter can be related to the theoretical framework developed by Garicano and Rossi-Hansberg (2006) and Garicano and Hubbard (2007), who model firms as knowledge-based hierarchies. According to this strand of literature, knowledge is a typical intangible asset which is complementary to physical inputs involved in vertically linked products. Knowledge can be accumulated for example by hiring better managers and adopting better managerial procedures, but it has a fixed cost. Therefore, given a firm size (in our case group size), best intangible assets can be shared in the presence of a larger number of units of production and hence their costs can be smoothed on a larger scale. Our results add to these findings, by showing that, within a Business Group structure, besides the positive relationship between group size and performance, also the complexity of the command chain is positively correlated with productivity.

In an exercise not reported here we included as a further control the simple number of group affiliates, which turned out to be not significant and quasi-collinear to group employment. At the same time the magnitude and significance of the coefficient for hierarchical complexity was practically unaltered. The latter result points at the fact that group employment and number of affiliates can be replaceable proxies for a group size, but not for a group organization.³²

Our results also show that too complex an organization can also be problematic. As previously discussed, our measure of hierarchical complexity should proxy the higher fixed costs of communication between the single affiliate and the headquarter when more than one layer of management is involved in the decision process. From the point of view of the headquarter, it might be efficient to decentralize decisions at lower levels of hierarchies to better exploit intangible assets, but at the same time, as coordinating a decision with a distant affiliate is relatively more cumbersome, internal coordination costs can become so high that the organization becomes too complex to be managed efficiently. To

³¹Besides the possible omitted variable bias due to the exclusion of a measure of hierarchical complexity, Atalay, Hortacsu and Syverson (2012) limited their analysis to plants within firms, without an explicit control for ownership structures developed by Business Groups. As discussed, this latter feature might lead to an underestimation of the level of vertical integration actually occurring within some of their non-independent firms.

³²Reciprocal sample properties for number of affiliates (N) and the Global Index of Complexity (GIC) are reported in Appendix C.

capture this effect, in column 5 of Table 9 we have introduced a squared term in our hierarchical complexity variable, which turns out to be negative and significant. This latter result is 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.

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.³³ Such an evidence of marginally decreasing returns from increasing complexity is relevant, as it puts a natural limit to the growth in complexity: 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 assumption of increasing marginal costs of communication across hierarchical levels, and hence the robustness of a non-linear relation between hierarchical complexity and productivity, is further tested in the last specification of Table 9. There we have introduced as an additional robustness check a control for the simple hierarchical distance of the affiliate, i.e. the length of the command chain linking each affiliate to the parent company. We find that on average the further the firm is from the decision making center, the lower its level of productivity appears to be. But the latter result only holds when we control at the same time for our measure of hierarchical complexity. When considering only hierarchical distance in the model, i.e. excluding hierarchical complexity, affiliates located at further levels of control would actually display higher levels of productivity.

The latter is a refinement of the delegation view of corporate group structure: on the one hand, as already discussed, sequential delegation of authority allows for the exploitation of economies of scale for knowledge inputs, and hence affiliates belonging to bigger (more complex) networks, no matter how this is proxied (that is, hierarchical distance or hierarchical complexity), are relatively more productive.³⁴ At the same time, the finding that subsidiaries located at further hierarchical distances from the headquarters show (at the margin) a negative productivity premium once controlling for the overall hierarchical complexity of the group shows that additional layers may give rise to extra costs.

5 Conclusions

In this paper we have exploited a unique global dataset on Business Groups in order to explore the relation between institutional characteristics of host countries, organizational choices of Business Groups and performance, as proxied by productivity of their affiliates.

To that extent, we have improved upon traditional measures of vertical integration, distinguishing

³³Note that our Group Index of Complexity is not strictly monotonous in either number of affiliates or levels of distance, and thus a given value can identify a range of cases with different combinations in terms of numbers of affiliates or control levels.

³⁴Assuming that a higher number of layers of management (both in levels and nodes) could imply a more complex /delegated decision process, our positive correlation between hierarchical complexity and productivity is also consistent with the findings by Bloom, Sadun and Van Reenen (2012), according to which more delegation of authority implies a higher firm-level performance thanks to a better reallocation of resources.

between an affiliate- and a group-level propensity to exchange intermediate inputs. We have also discussed why, studying Business Groups, it is necessary to complement these measures with a novel metric able to account for the hierarchical complexity of the command chain linking affiliates and headquarters. Our novel metric borrows from graph theory and is consistent with a refinement of the delegation view of corporate group structure modelling firms as knowledge-based hierarchies.

We have shown how the decision process related to the design of optimal organization of Business Groups is multi-dimensional, as vertical integration choices depend on host countries' characteristics, as acknowledged by the literature, but the latter also influence decisions on the hierarchical organization of the group, which in turn affects the same vertical integration choices. Thus, ignoring both the impending ownership structure that links firms in a common hierarchy and the interplay that occurs between integration and organization might lead to unobserved omitted variable biases, for example when studying the institutional drivers of production networks.

Our results also point at a positive and significant robust relationship between affiliates' productivity and hierarchical complexity which dominates the positive relationship between performance and vertical integration. This finding is consistent with the idea that knowledge is a typical intangible asset complementary to physical inputs involved in vertically linked products, such that more complex structures allow for a smoothing of fixed costs thanks to a well designed mechanism of transmission along the hierarchy. Still, we find the latter relationship to be partly non-linear, reflecting a trade-off between the accumulation of knowledge and its communication down the hierarchy.

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 hierarchical 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.

A third line of research is related to the exploration of the role of internal capital markets developed by Business Groups for the allocation of financial resources among competing investment projects. Some preliminary evidence we have obtained on our data shows that the possibility to shift financial funds across activities and across countries mitigates the financial pressures exerted on corporate structures with respect to stand-alone firms, in line with previous findings of the finance literature. At the same time, however, we detect a complementarity across BGs between internal capital markets and an enhanced borrowing capacity, probably due to a better reputation that acts as an intangible collateral on credit markets. The latter seems to be particularly relevant in the years of financial crisis and as such constitutes a potential channel of competitiveness that is worth further investigation.

References

- Acemoglu, Daron, Pol Antràs, and Elhanan Helpman, "Contracts and Technology Adoption", *American Economic Review*, 97(2007), 916-943.
- Acemoglu, Daron, Simon Johnson, and Todd Mitton, "Determinants of Vertical Integration: Financial Development and Contracting Costs". *The Journal of Finance*, 63(2009), 1251-1290.
- Acharya, Viral V., Stewart G. Myers, and Raghuram G. Rajan, "The Internal Governance of Firms", *The Journal of Finance*, 66(2011), 689-720.
- Aghion, Philippe, and Richard Holden, "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?." *Journal of Economic Perspectives*, 25(2011): 181-197.
- Alfaro, Laura, and Andrew Charlton, "Intra-industry Foreign Direct Investment", *American Economic Review*, 99(2009), 2096-2119.
- Alfaro, Laura, Paola Conconi, Harald Fadinger, and Andrew Newman, "Trade Policy and Firm Boundaries", NBER Working Paper N. 16118, 2011.
- Almeida, Heitor V., and Daniel Wolfenzon, "A Theory of Pyramidal Ownership and Family Business Groups", *The Journal of Finance*, 61(2006), 2637-2680.
- Altomonte, Carlo, Filippo di Mauro, Gianmarco I. P. Ottaviano, Armando Rungi, and Vincent Vicard, "Global Value Chains during the Great Trade Collapse: a Bullwhip Effect?", in S. Beugelsdijk, S. Brakman, H. van Ees, H. Garretsen (eds.) *Firms in the International Economy*, MIT Press, Cambridge MA., 2013.
- Antràs, Pol, "Firms, Contracts, and Trade Structure", *The Quarterly Journal of Economics*, 118(2003), 1375-1418.
- Antràs, Pol, and Elhanan Helpman. "Global Sourcing," *Journal of Political Economy*, 112(2004), pp. 552-580.
- , "Contractual Frictions and Global Sourcing," in *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. (Cambridge, MA: Harvard University Press, 2008).
- Antràs, Pol and Esteban Rossi-Hansberg, "Organizations and Trade," *Annual Review of Economics*, 1(2009), 43-64.
- Atalay, Enghin, Ali Hortaçsu, and Chad Syverson. "Why Do Firms Own Production Chains?", NBER Working Paper N. 18020, 2012.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. "Bringing the Market Inside the Firm?", *American Economic Review Papers and Proceedings*, 91 (2001), 212-218.
- , "Relational Contracts and the Theory of the Firm", *The Quarterly Journal of Economics*, 117 (2002), pp. 39-84
- Beck, Thorsten, and Asli Demirgüç-Kunt, "Financial Institutions and Markets Across Countries and over Time: Data and Analysis", World Bank Policy Research Working Paper No. 4943, 2009.
- Beirlant, Jan, Petra Vynckier and Jozef L. Teugels, "Excess Functions and Estimation of Extreme-Value Index", *Bernoulli* 2(1996), 293-318
- Belenzon, 2012
- Belenzon, 2013a
- Belenzon, 2013b
- Bloom, Nicholas, Benn Eifert, David McKenzie, Aprajit Mahajan, and John Roberts, "Does management matter: evidence from India", *The Quarterly Journal of Economics*, 128(2013), 1-51.
- Bloom, Nicholas, Raffaella Sadun and John Van Reenen, "Does Product Market Competition Lead Firms to Decentralize?", *American Economic Review Papers and Proceedings*, 100(2010), 434-438.

- , "Americans do I.T. Better: US Multinationals and the Productivity Miracle", *American Economic Review*, 102(2012), 167-201.
- , "The Organization of Firms across Countries", *The Quarterly Journal of Economics*, 127(2012), 1663-1705.
- Caliendo, Lorenzo, and Esteban Rossi-Hansberg (2012). "The Impact of Trade on Organization and Productivity", *The Quarterly Journal of Economics*, 127(2012), 1393-1467.
- Coles, Stuart, *An Introduction to Statistical Modelling of Extreme Values*. Springer, 2001.
- Colpan, Asli, and Takashi Hikino, "Foundations of Business Groups: Towards an Integrated Framework" in *The Oxford Handbook of Business Groups*, Colpan et al. (eds). Oxford University Press, 2010.
- Emmert-Streib, Frank, Matthias Dehmer, "Information Theoretic Measures of UHG Graphs with Low Computational Complexity," *Applied Mathematics and Computation*, 190(2007), 1783-1794.
- Eurostat (2007). Recommendations Manual on the Production of Foreign Affiliates Statistics (FATS). European Commission.
- Fan, Joseph, and Larry Lang, "The measurement of relatedness: An application to corporate diversification", *Journal of Business*, 73(2000), 629-660.
- Fruin W. Mark, "Business Groups and Inter-firm Networks" in *The Oxford Handbook of Business History*, Geoffrey J. and Zeitlin J. (eds). Oxford: Oxford University Press, 2008.
- Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production", *Journal of Political Economy*, 108(2000), 874-904.
- Garicano, Luis, and Thomas N. Hubbard, "The Return to Knowledge Hierarchies.", NBER Working Paper N. 12815, 2007.
- Garicano, Luis, and Esteban Rossi-Hansberg, "Inequality and the Organization of Knowledge," *American Economic Review*, 94(2004), 197-202.
- , "Organization and Inequality in a Knowledge Economy.", *The Quarterly Journal of Economics*, 121(2006), 1383-1435.
- , "Organizing Growth", *Journal of Economic Theory*, 147(2012), 593-629.
- Granovetter, Mark, "Coase revisited: Business Groups in the Modern Economy", *Industrial and Corporate Change* 4(1995), 93-130.
- Grossman, Sanford, and Oliver Hart, "The costs and benefits of ownership: A theory of vertical and lateral Integration", *Journal of Political Economy* 94(1986), 691-719.
- Grossman, Gene M., and Elhanan Helpman, "Integration versus Outsourcing in Industry Equilibrium", *The Quarterly Journal of Economics* 117(2002), 85-120.
- , "Outsourcing versus FDI in Industry Equilibrium", *Journal of the European Economic Association*, 1(2003), 317-327.
- , "Managerial incentives and the international organization of production", *Journal of International Economics*, 63(2004), pp. 237-262.
- , "Outsourcing in a Global Economy," *Review of Economic Studies*, 72(2005), 135-159.
- Hart, Oliver, and Bengt Holmstrom, "A Theory of Firm Scope", *The Quarterly Journal of Economics*, 125 (2010), 483-513.
- Hart, Oliver, and John Moore, "Property rights and the nature of the firm", *Journal of Political Economy*, 98(1990), 1119-1158.
- Helpman, Elhanan, "Trade, FDI, and the Organization of Firms", *Journal of Economic Literature*, 44(2008), 589-630.

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

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

IRI (2011). *The 2010 EU Industrial R&D Investment SCOREBOARD*. European Commission, Bruxelles (available at <http://iri.jrc.ec.europa.eu/docs/scoreboard/2010/SB2010.pdf>).

Jones, Geoffrey, and Asli Colpan, "Business Groups in Historical Perspective" in *The Oxford Handbook of Business Groups*, Colpan et al. (eds.). Oxford: Oxford University Press, 2010.

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

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

–, "Business Groups, the Financial Market and Modernization", *Economics of Transition*, 11(2003), 671-696.

Khanna and Palepu, 1997

Khanna, Tarun, and Yishai Yafeh, "Business Groups in Emerging Markets: Paragons or Parasites?", *Journal of Economic Literature*, 45(2007), 331-372.

Kumar, Krishna B., Raghuram G. Rajan, and Luigi Zingales, "What determines firm size?", NBER Working Paper N. 7208, 2001.

Lafontaine, Francine, and Margaret Slade, "Vertical Integration and Firm Boundaries: The Evidence," *Journal of Economic Literature*, 45(2007), 629-685.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Schleifer, "Corporate Ownership around the World", *Journal of Finance*, 54(1999), 471-517.

Leff

Maksimovic, Vojislav, Gordon Phillips, and N.R. Prabhala, "Post-Merger Restructuring and the Boundaries of the Firm." *Journal of Financial Economics*, 102(2011), 317-343.

Marin, Dalia, and Thierry Verdier, "Power in the Multinational Corporation in Industry Equilibrium," *Economic Theory*, 38(2009), 437-464.

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

Nunn, Nathan, and Daniel Treffer, "The Boundaries of the Multinational Firm: An Empirical Analysis", in *The Organization of Firms in a Global Economy*, E. Helpman, D. Marin, and T. Verdier (eds.), Harvard University Press, 2008.

OECD (2005). *Guidelines for Multinational Enterprises*. OECD, Paris.

Palacz, Wojciech, "Algebraic Hierarchical Graph Transformation", *Journal of Computer and System Sciences*, 68(2004), 497-520.

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

Rajan, Raghuram G., and Luigi Zingales, "Financial dependence and growth", *American Economic Review*, 88(1998), 559-586.

–, "The Firm As a Dedicated Hierarchy: A Theory Of The Origins And Growth Of Firms", *The Quarterly Journal of Economics*, 116(2001), 805-851.

--, "The Influence of the Financial Revolution on the Nature of Firms," *American Economic Review*, 91(2001), 206-211.

Schoar, Antoinette, "The Effect of Diversification on Firm Productivity." *Journal of Finance*, 62(2002), 2379-2403.

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, Geneva.

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*. US Bureau of Economic Analysis, November 2012 (available at: <http://www.bea.gov/scb/pdf/2012/11%20November/1112MNCs.pdf>).

Vitali, Stefania, James B. Glattfelder, and Stefano Battiston, "The Network of Global Corporate Control". *PLoS ONE* 6(2011): e25995. doi:10.1371/journal.pone.0025995.

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

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

--, *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press, 1975

--, *The Economic Institutions of Capitalism*. New York: Free Press, 1985.

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

Our two main sources of data are both compiled by Bureau Van Dijk (BvD), a Belgian 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 A.1 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 Dijk 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 A.1: Original sources of ownership linkages collected by Bureau Van Dijk

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), Factsset, 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), Kredidiinfo (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. In case no consolidated accounts are available we include the affiliate in the group where it is located at shorter control distance from the parent.

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 preliminarily in two groups:

- a) a set of independent companies, that have as controlling shareholder individuals or a family or no specific corporate 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 A.2 for a description of this latter sample).

The algorithm then screens 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.

³⁵See Vitali, Glattfelder and Battistoni (2011) for an assessment of existing methodologies to attribute global corporate control and their limits.

³⁶If 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.

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.

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 closest level where we have finally 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 at least one financial account that we use for the calculation of premia in Table 2.

In Table A.2 we report some descriptive statistics of the control group collecting non-affiliated firms confronted with the set of affiliates belonging to Business Groups.

Table A.2: A control group of non-affiliated firms, some descriptive statistics (US thousand dollars)

Variables	Non-affiliated firms*						Affiliated firms					
	N. observations	Mean	Median	St. dev.	Min	Max	N. observations	Mean	Median	St. dev.	Min	Max
Employment	2,952,128	13	4	51	1	14717	1,005,381	367	25	6225	1	1917456
Fixed assets	3,756,003	931	70	15016	1	4593066	1,005,381	109404	728	1812076	1	2.79E+08
Turnover (sales)	3,756,003	1464	279	4344	1	979128	1,005,381	119045	3493	1616124	1	3.57E+08

*Affiliated firms are independent firms: controlling shareholders are individuals or families or no specific corporate entity; they do not control affiliates.

Appendix B:

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

Table B1: Correspondence table NAICS 2002 and US Input Output tables, source: US Bureau
of Economic Analysis

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
Federal Reserve banks, credit intermediation, and 41 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
Rental and leasing services and lessors of intangible 46 assets	532RL	532, 533
47 Legal services	5411	5411
48 Computer systems design and related services	5415	5415
Miscellaneous professional, scientific, and technical 49 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
Performing arts, spectator sports, museums, and 57 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 3.2 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 headquarters) 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, in order to discount a high communication costs with the decision-making centre. Hence, 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 sample of Business Groups.

In Table C1 we report some descriptive statistics that already show how both the node entropy in the second column and the *GIC* in the third column 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 has a right tail from the 75th percentile.

Table C.1: Comparison of distributions: number of affiliates (N), node entropy ($H(G)$) 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
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 *GIC* can be computed. 13 Business Groups are excluded from computation for lack of information on hierarchical levels.

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 logarithmic weight of formulas 3 and 4, also groups that have few affiliates but all positioned on a same control level ($N = n_l$, whatever l) end up with a $H(G)$ and a *GIC* that both have null figures until the 83rd percentile.

The latter 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 C.2 we show however how sample distributions of *GIC* and N are significantly correlated when we report descriptive statistics of the first by Business Groups' size classes

Table C.2: Descriptives of *GIC* by size classes of Business Groups

N. affiliates	Frequency	mean <i>GIC</i>	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 size class there is a certain (increasing) degree of heterogeneity of hierarchical design, confirming that the *GIC* is more able to catch that heterogeneity. Indeed, in the last column, where we report maximum *GIC* by size classes, 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 levels³⁷.

³⁷More precisely, given N number of affiliates and L number of levels for each Business Group, *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 m th and n th level. That is, when the group/graph is pyramidal.

One would like to graphically compare N and GIC to observe how they behave along our sample distributions. Given their extremely long right tails, we resort to mean excess plots of Figure C.1, where graphs are the results of the following so called excess mean function (see Beirlant, Vynckier and Teugels, 1996; Coles, 2001):

$$e_F(u) = E(Y - u | Y > u) \tag{C1}$$

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

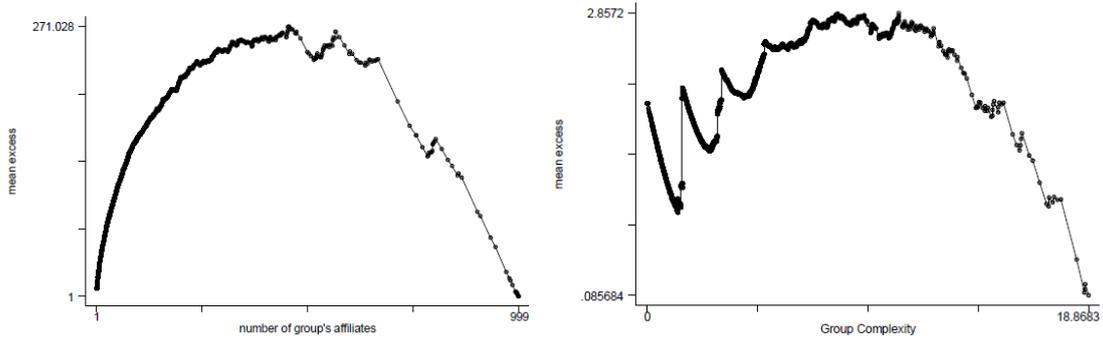
Shortly, the mean excess plot gives back a transformation of the distribution such that it is represented as the excess of the consecutive sample means of a random variable Y with respect to a sequence of threshold values u picked from the same domain of random variable Y such that the cumulative distribution $F(x)$ is defined on the interval $(0, 1)$. More analytically, 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 C.1 begin with the first threshold value u different from zero as present in our sample of Business Groups (1 in the case of N and .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 C.1: A visual comparison of number of affiliates (N) and GIC distributions: mean excess functions

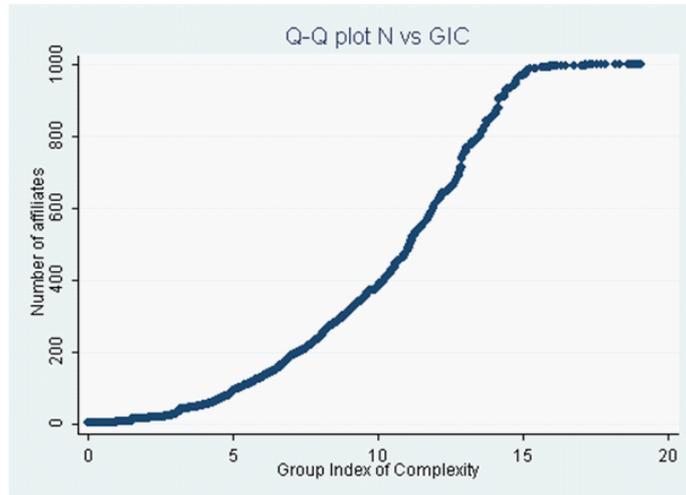


In both cases, we observe that after the maximum of the mean excess function, the distributions are steeply decreasing almost linearly. It means that in original (non-transformed) distributions they both report a long right-tail that is very much similar from a statistical point of view. The observation

of a steeply decreasing right tail after a mean excess transformation is usually considered itself as a visual test to assume that the original distribution can be proxied by a Pareto. However, on the left tail, number of affiliates and GIC behave somehow differently since some points of discontinuity are present for the *GIC* excess function but not for the *N*'s excess function. Those discontinuities reveal that at some lower measures of hierarchical complexity there is a high density of Business Groups, a sample feature that was not observed when looking at Tables C.1 and C.2.

Finally, in Figure C.2 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 the *GIC* has a smoother distribution especially in the last quartile of the distribution. Thus, the *GIC* is able to provide a continuous measure for organization of command chains that differentiates more when groups have a higher number of affiliates.

Figure C.2: A quantile-quantile plot of number of affiliates (*N*) vs GIC



References

- Acemoglu, Daron, Pol Antràs, and Elhanan Helpman, "Contracts and Technology Adoption", *American Economic Review*, 97(2007), 916-943.
- Acemoglu, Daron, Simon Johnson, and Todd Mitton, "Determinants of Vertical Integration: Financial Development and Contracting Costs". *The Journal of Finance*, 63(2009), 1251-1290.
- Acharya, Viral V., Stewart G. Myers, and Raghuram G. Rajan, "The Internal Governance of Firms", *The Journal of Finance*, 66(2011), 689-720.
- Aghion, Philippe, and Richard Holden, "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?." *Journal of Economic Perspectives*, 25(2011): 181-197.
- Alfaro, Laura, and Andrew Charlton, "Intra-industry Foreign Direct Investment", *American Economic Review*, 99(2009), 2096-2119.
- Almeida, Heitor V., and Daniel Wolfenzon, "A Theory of Pyramidal Ownership and Family Business Groups", *The Journal of Finance*, 61(2006), 2637-2680.
- Antràs, Pol, "Firms, Contracts, and Trade Structure", *The Quarterly Journal of Economics*, 118(2003), 1375-1418.
- Antràs, Pol, and Elhanan Helpman. "Global Sourcing," *Journal of Political Economy*, 112(2004), pp. 552–580.
- –, "Contractual Frictions and Global Sourcing," in *The Organization of Firms in a Global Economy*, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. (Cambridge, MA: Harvard University Press, 2008).
- Antràs, Pol and Esteban Rossi-Hansberg, "Organizations and Trade," *Annual Review of Economics*, 1(2009), 43-64.
- Atalay, Englin, Ali Hortaçsu, and Chad Syverson. "Why Do Firms Own Production Chains?", NBER Working Paper N. 18020, 2012.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. "Bringing the Market Inside the Firm?", *American Economic Review Papers and Proceedings*, 91 (2001), 212-218.
- –, "Relational Contracts and the Theory of the Firm", *The Quarterly Journal of Economics*, 117 (2002), pp. 39-84
- Beck, Thorsten, and Asli Demirgüç-Kunt, "Financial Institutions and Markets Across Countries and over Time: Data and Analysis", World Bank Policy Research Working Paper No. 4943, 2009.
- Beirlant, Jan, Petra Vynckier and Jozef L. Teugels, "Excess Functions and Estimation of Extreme-Value Index", *Bernoulli* 2(1996), 293-318
- Belenzon, Sharon, Tomer Berkovitz, and Luis A. Rioz. "Capital Markets and Firm Organization: How Financial Development Shapes European Corporate Groups", *Management Science*, 59(2013a), pp. 1326-1343.
- Belenzon, 2013b
- Bloom, Nicholas, Benn Eifert, David McKenzie, Aprajit Mahajan, and John Roberts, "Does management matter: evidence from India", *The Quarterly Journal of Economics*, 128(2013), 1-51.
- Bloom, Nicholas, Raffaella Sadun and John Van Reenen, "Does Product Market Competition Lead Firms to Decentralize?", *American Economic Review Papers and Proceedings*, 100(2010), 434-438.
- –, "Americans do I.T. Better: US Multinationals and the Productivity Miracle", *American Economic Review*, 102(2012), 167-201.
- –, "The Organization of Firms across Countries", *The Quarterly Journal of Economics*, 127(2012), 1663-1705.
- Caliendo, Lorenzo, and Esteban Rossi-Hansberg (2012). "The Impact of Trade on Organization and Productivity", *The Quarterly Journal of Economics*, 127(2012), 1393-1467.
- Coles, Stuart, *An Introduction to Statistical Modelling of Extreme Values*. Springer, 2001.

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

Emmert-Streib, Frank, Matthias Dehmer, "Information Theoretic Measures of UHG Graphs with Low Computational Complexity," *Applied Mathematics and Computation*, 190(2007), 1783-1794.

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

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

Fruin W. Mark, "Business Groups and Inter-firm Networks" in *The Oxford Handbook of Business History*, Geoffrey J. and Zeitlin J. (eds). Oxford: Oxford University Press, 2008.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production", *Journal of Political Economy*, 108(2000), 874-904.

Garicano, Luis, and Thomas N. Hubbard, "The Return to Knowledge Hierarchies.", NBER Working Paper N. 12815, 2007.

Garicano, Luis, and Esteban Rossi-Hansberg, "Inequality and the Organization of Knowledge," *American Economic Review*, 94(2004), 197–202.

– –, "Organization and Inequality in a Knowledge Economy.", *The Quarterly Journal of Economics*, 121(2006), 1383-1435.

– –, "Organizing Growth", *Journal of Economic Theory*, 147(2012), 593–629.

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

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

Grossman, Gene M., and Elhanan Helpman, "Integration versus Outsourcing in Industry Equilibrium", *The Quarterly Journal of Economics* 117(2002), 85-120.

– –, "Outsourcing versus FDI in Industry Equilibrium", *Journal of the European Economic Association*, 1(2003), 317-327.

– –, "Managerial incentives and the international organization of production", *Journal of International Economics*, 63(2004), pp. 237-262.

– –, "Outsourcing in a Global Economy," *Review of Economic Studies*, 72(2005), 135-159.

Hart, Oliver, and Bengt Holmstrom, "A Theory of Firm Scope", *The Quarterly Journal of Economics*, 125 (2010), 483-513.

Hart, Oliver, and John Moore, "Property rights and the nature of the firm", *Journal of Political Economy*, 98(1990), 1119–1158.

Helpman, Elhanan, "Trade, FDI, and the Organization of Firms", *Journal of Economic Literature*, 44(2008), 589-630.

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

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

IRI (2011). *The 2010 EU Industrial R&D Investment SCOREBOARD*. European Commission, Bruxelles (available at <http://iri.jrc.ec.europa.eu/docs/scoreboard/2010/SB2010.pdf>).

Jones, Geoffrey, and Asli Colpan, "Business Groups in Historical Perspective" in *The Oxford Handbook of Business Groups*, Colpan et al. (eds.). Oxford: Oxford University Press, 2010.

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

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

—, "Business Groups, the Financial Market and Modernization", *Economics of Transition*, 11(2003), 671-696.

Khanna and Palepu, 1997

Khanna, Tarun, and Yishai Yafeh, "Business Groups in Emerging Markets: Paragons or Parasites?", *Journal of Economic Literature*, 45(2007), 331-372.

Kumar, Krishna B., Raghuram G. Rajan, and Luigi Zingales, "What determines firm size?", NBER Working Paper N. 7208, 2001.

Lafontaine, Francine, and Margaret Slade, "Vertical Integration and Firm Boundaries: The Evidence," *Journal of Economic Literature*, 45(2007), 629-685.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Schleifer, "Corporate Ownership around the World", *Journal of Finance*, 54(1999), 471-517.

Leff

Maksimovic, Vojislav, Gordon Phillips, and N.R. Prabhala, "Post-Merger Restructuring and the Boundaries of the Firm." *Journal of Financial Economics*, 102(2011), 317-343.

Marin, Dalia, and Thierry Verdier, "Power in the Multinational Corporation in Industry Equilibrium," *Economic Theory*, 38(2009), 437-464.

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

Nunn, Nathan, and Daniel Treffer, "The Boundaries of the Multinational Firm: An Empirical Analysis", in *The Organization of Firms in a Global Economy*, E. Helpman, D. Marin, and T. Verdier (eds.), Harvard University Press, 2008.

OECD (2005). *Guidelines for Multinational Enterprises*. OECD, Paris.

Palacz, Wojciech, "Algebraic Hierarchical Graph Transformation", *Journal of Computer and System Sciences*, 68(2004), 497-520.

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

Rajan, Raghuram G., and Luigi Zingales, "Financial dependence and growth", *American Economic Review*, 88(1998), 559-586.

—, "The Firm As a Dedicated Hierarchy: A Theory Of The Origins And Growth Of Firms", *The Quarterly Journal of Economics*, 116(2001), 805-851.

—, "The Influence of the Financial Revolution on the Nature of Firms," *American Economic Review*, 91(2001), 206-211.

Schoar, Antoinette, "The Effect of Diversification on Firm Productivity." *Journal of Finance*, 62(2002), 2379-2403.

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, Geneva.

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*. US Bureau of Economic Analysis, November 2012 (available at: <http://www.bea.gov/scb/pdf/2012/11%20November/1112MNCs.pdf>).

Vitali, Stefania, James B. Glattfelder, and Stefano Battiston, "The Network of Global Corporate Control". *PLoS ONE* 6(2011): e25995. doi:10.1371/journal.pone.0025995.

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

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

–, *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press, 1975

–, *The Economic Institutions of Capitalism*. New York: Free Press, 1985.