The Journal of Law, Economics, and Organization, 2026, **00**, 1–50
https://doi.org/10.1093/jleo/ewag011
Article

OXFORD

Business groups as knowledge-based hierarchies of firms

Carlo Altomonte[1,2,*], Gianmarco I. P. Ottaviano[2,3,4], Armando Rungi[5], Tommaso Sonno[2,4,6]

[1]Department of Social and Political Sciences, Bocconi University, Milan, Italy
[2]Baffi Carefin, Bocconi University, Milan, Italy
[3]Department of Economics, Bocconi University, Milan, Italy
[4]Centre for Economic Performance, London School of Economics, London, United Kingdom
[5]Firms Markets Technologies Lab, IMT School for Advanced Studies, Lucca, Italy
[6]Department of Economics, University of Bologna, Bologna, Italy

*Corresponding author: carlo.altomonte@unibocconi.it

ABSTRACT

We provide the first worldwide overview of the patterns of hierarchical differentiation across Business Groups (BGs), highlighting the coexistence of different hierarchical shapes. We show how the different shapes can arise as optimal hierarchical structures in a knowledge-based model of BGs when subsidiaries' operations involve problem-solving under parents' supervision. The optimal choice of hierarchical structure is driven by production efficiency and two dimensions of problem-solving: efficiency related to supervising knowledge creation and handling associated communication across subsidiaries. We check the consistency of the model's predictions with the empirical patterns. The model successfully passes the consistency test (*JEL* D23, L23, F23, L25, G34)

1. INTRODUCTION

A Business Group (BG) is an organizational form of economic activity in which at least two legally autonomous firms function as a single economic entity through hierarchical control: a parent company ("headquarters" [HQ]) owns, directly or indirectly, the majority of the equity shares of at least one legally independent firm ("subsidiary"). Under this definition, the world's largest corporations by consolidated revenue, as classified in the Fortune 500 list, are all organized as BGs. According to our data, in 2022, more than half of the multinational value added is accounted for by about 1% BGs with more than 100 subsidiaries located in different countries. Data from the US Bureau of Economic Analysis also show that at least 75% of total US trade can be linked to firms operating in the USA as parts of BGs (either as US HQs or as US subsidiaries of foreign groups). In the case of France, around 65% of aggregate imports or exports can be attributed to firms that belong to BGs (Altomonte and Rungi 2013).

BGs vary significantly in terms of organization. In our data, they range from very simple structures, with one HQ controlling only one subsidiary in a hierarchical layer below it, to large multinational groups featuring more than 670 subsidiaries organized across more than 10 hierarchical layers. Moreover, even BGs with very similar portfolios of activities may exhibit very different hierarchical structures. This is the case, for example, of two well-known BGs, General Motors and Mitsubishi, both operating in the automotive industry. In our data, the two BGs are comparable in size: 685 subsidiaries in 55 countries for General Motors and 528 subsidiaries in 47 countries for Mitsubishi. However, General Motors is organized in 10 hierarchical layers covering 26 sectors (three-digit NAICS 2002), while Mitsubishi has half the number of hierarchical layers (five) but spans almost twice the number of sectors (44).

However, even across very different hierarchical structures, the position of subsidiaries in the hierarchy seems to follow some regular patterns, depending on the nature of their activities with respect to the parent company. For instance, when Google reorganized its BG, it created a new HQ, Alphabet, directly controlling Google Inc., which in turn kept control of most of the other subsidiaries of the group. A few subsidiaries, however, ended up being placed under the direct control of Alphabet and Google Inc. Among those were Calico, a biotech company researching the mechanisms of ageing, and Crystal Computing, a new data center to improve European users' access to Google databases. Both activities are less standard with respect to the core products of Google Inc., a pattern of assignment common across BGs in our data.

Despite vast heterogeneity in organization, little is known about the drivers of hierarchical differentiation across BGs, or about the reasons behind the positioning of specific subsidiaries within the hierarchy. To shed more light on these issues, we have built a cross-section of some 2.9M parent companies controlling 5.7M subsidiaries worldwide based on the Orbis Ownership database. We have merged these data with information on firms' patents and citations, as retrieved from the Orbis Intellectual Property Database, thus constructing a network of patent-citation links (1.4 million patent applications and a network of 9.5 million firm-patent citations). Descriptive statistics reveal three novel stylized facts. First, there is a continuum of different hierarchical shapes of BGs, from pyramidal-shaped hierarchies (where subsidiaries are less dense in layers closer to the HQ) to inverted ones (where subsidiaries are denser in layers closer to the HQ). This finding contrasts with the prevailing idea that pyramidal hierarchies are the norm in the organization of BGs, as in the studies on formal hierarchical structures and team performance (e.g., in Wellman et al. [2020]). Second, as the hierarchical distance of subsidiaries from the HQ increases, their activities tend to become more standard and easier to routinize.[1] Third, within BGs, there is a systematic concentration of innovative activities in the HQ or in the subsidiaries placed in layers close to HQ as measured in terms of patents and citations. Moreover, layers that are more distant from the HQ are more likely to cite its patents or patents of subsidiaries in layers closer to it than vice versa. This also holds after excluding citations between subsidiaries in adjacent layers, suggesting that knowledge flows between subsidiaries within BGs are not necessarily associated with direct ownership links. This new set of robust stylized facts constitutes the first contribution of our article to the literature.

The second contribution of the article is the rationalization of those facts through a parsimonious theoretical framework based on the idea that knowledge is an essential input for

[1] Routinizability is measured following Blinder and Krueger (2009, 2013)

BGs' operations regardless of their industries. In particular, the article proposes a model in which the hierarchical structure of a BG is designed by the HQ to efficiently delegate to subsidiaries knowledge creation, diffusion, and application to production when its direct involvement is time-constrained. For the subsidiaries to produce, several problems of different importance and corresponding difficulty have to be solved. The HQ knows how to solve each problem but has limited time to do that. It, therefore, allocates problems of different difficulty to subsidiaries according to their ability to solve them. However, regardless of their abilities, subsidiaries still need the HQ's advice to succeed in solving the assigned problems. Advice to a subsidiary of a given ability can be direct or indirect through subsidiaries of a higher ability. As the former is more time-consuming, the HQ prefers to directly advise higher-ability subsidiaries tackling more difficult and important problems and indirectly advise the other subsidiaries. The BG's hierarchical structure emerges as an optimal arrangement for creating, diffusing, and applying proprietary knowledge, with its shape depending on the interactions among three primitive characteristics of the group: production efficiency and two dimensions of problem-solving efficiency, namely efficiency in advising knowledge creation and efficiency in handling associated communication across subsidiaries. BGs with higher production efficiency have more layers and subsidiaries. Pyramids and inverted pyramids are chosen by BGs with higher and lower problem-solving efficiency, respectively. Intuitively, for BGs with lower efficiency in advising and communicating, it is better to keep problem-solving, and thus production, on layers closer to the parent. Moreover, consistent with what is observed in the data, the model predicts that the HQ delegates tasks to different layers in descending order of difficulty from the top of the hierarchy.

In a quantitative exercise, we leverage the model's equilibrium equations to estimate the parameters regulating problem difficulty and the two dimensions of problem-solving efficiency. Using these estimates, along with the model's structure and other estimated parameters from the literature, we calibrate the parameter regulating production efficiency. We then use the calibrated equations to conduct various simulations and show that the model correctly replicates our main stylized facts. In particular, the model accurately predicts that BGs organized in a larger number of hierarchical layers tend to have more subsidiaries and are less likely to be organized as pyramids. It also correctly predicts the cross-country variation in the frequency of pyramids in our data. As a sanity check for the parameter values of problem-solving efficiency estimated through the model, we compare their cross-country variation with that extracted from related questions in the World Management Survey, obtaining similar results. This establishes a link between managerial practices and the organization of BGs, which represents the third contribution of our article to the literature.

Scholars from different fields of study have motivated the emergence of BGs along three main dimensions of resource allocation across their subsidiaries related to capital, labor and taxable income. A number of scholars have shown how resources can be moved across subsidiaries to fund those with higher return on investment or bail out those in financial distress (Khanna and Palepu 2000; Khanna and Rivkin 2001; Almeida and Wolfenzon 2006; Almeida et al. 2011; Belenzon and Schankerman 2013). Evidence also exists on the existence of labor markets internal to BGs, whereby workers are reallocated across subsidiaries in response to demand shocks (Huneeus et al. 2021; Cestone et al. 2023). Lastly, BGs have been shown to shift taxable income across subsidiaries for tax arbitrage purposes (Lewellen and Robinson 2013).

Our analysis introduces knowledge as a fourth critical dimension of resource allocation across subsidiaries within BGs. This aligns with seminal work by Bartlett and Ghoshal (1989), according to whom (p. 197) for large transnational organizations, "the most difficult task is to coordinate the voluminous flow of strategic information and proprietary knowledge required to operate them." As "the sheer volume and complexity of information would overload HQs if coordination were centralized," the most effective way to achieve coordination is "by transferring personnel with the relevant knowledge, or creating organizational forums that allow the free exchange of information and foster interunit learning." In fact, various types of knowledge flows have been documented, including "vertical" flows between parents and subsidiaries (Gupta and Govindarajan 1991; Bresman et al. 1999; Monteiro et al. 2008; Nair et al. 2015) as well as "horizontal" knowledge transfers between subsidiaries at the same hierarchical layer (Ciabuschi et al. 2011; Li and Lee 2015; Crespo et al. 2020; Garg et al. 2022; Wu et al. 2022). Our findings reveal a distinct pattern in the knowledge flows within a BG, whereby knowledge is predominantly created in the HQ or in closer subsidiaries engaging in more complex tasks, and subsequently diffused down the hierarchy.

In terms of data sources, we rely on the Orbis-Bureau van Dijk data, already used in the literature to study BGs, for instance with reference to innovation (Belenzon and Berkovitz 2010), the international transmission of shocks (Cravino and Levchenko 2017), or the effect of managerial culture on firm boundaries (Gorodnichenko et al. 2021). Other studies have used similar data to explore the boundaries of the firms belonging to BGs (e.g., Alfaro and Charlton 2009; Alfaro 2017) with respect to the BGs' external suppliers, or have started to exploit the Ownership Database of Orbis to map BGs through the notion of corporate control.[2] Belenzon et al. (2019) use Bureau Van Djik data on a subset of European BGs to investigate the relationship between BG structure and subsidiary autonomy, finding that the latter increases with the number of layers between the subsidiary and the HQ. More closely related to our work, Altomonte and Rungi (2013) use the notion of control to construct BGs' hierarchies in a smaller cross-section of groups across countries, deriving stylized facts on the positioning of subsidiaries along the hierarchy consistent with our analysis. Rungi et al. (2017) adopt a novel network framework to identify BG tree-like structures based on direct and indirect control, uncovering a strong concentration of corporate power in the hands of a few BGs, while they can have different configurations of ownership chains crossing multiple national borders. Sonno (2025) uses the notion of corporate control to create a worldwide panel of BGs' structures and uses these data to establish a causal link between multinational activities and episodes of conflict in developing countries. All these studies have made a number of methodological choices to construct BG's data, which are incorporated in our results and discussed in our robustness checks.[3]

The rest of the article is structured as follows. Section 2 presents the data sources and discusses the construction of the dataset. Section 3 presents the novel stylized facts derived from the dataset. Section 4 develops the knowledge-based model of BGs' hierarchical differentiation and checks the consistency of some of its key predictions with the empirical patterns. Section 5 concludes.

2. DATA AND DESCRIPTIVE STATISTICS

We define a BG as a collection of at least two legally autonomous firms functioning as a single economic entity through a common source of hierarchical control. Hierarchy in control implies that a parent company ("headquarters," or simply HQ) owns, partially or totally, the equity shares of a second legally independent firm. Moreover, the HQ may directly own the

[2] Some of the studies on external suppliers rely on data sourced from Dun&Bradstreet (D&B), which is one of the different sources now integrated in the Orbis Ownership database.

[3] Less closely related to this paper, but also using Orbis data to create a sample of BGs, Grosskurth (2019) exploits Orbis data to map the development of the global networks of multinational BGs, leveraging the number of controlled subsidiaries to identify a company's importance within a BG, but does not focus on the hierarchical shapes and their drivers; Aminadav and Papaioannou (2020) use ownership data to investigate ownership concentration and types of control across continents; Eppinger and Kukharskyy (2020) study domestic and cross-border ownership shares worldwide and find that firms have higher shares in subsidiaries located in countries with better contracting institutions.

equity shares of a third firm, thus placing it at the same (first) layer as the second firm in a shallower "flat" hierarchy. Alternatively, the second firm could control the third one, thus placing it at the second layer of a deeper "vertical" hierarchy. In this case, the HQ indirectly controls the third firm through the second. Multinational enterprises (MNEs) are a subset of BGs with at least one subsidiary abroad.[4]

To build our dataset on BGs worldwide, together with information on the organization of their internal hierarchical structure, we start with the Historical Ownership Database by Bureau Van Dijk. This dataset provides information on all of the company's shareholders, corporate or non-corporate (e.g., individuals or partnerships), and identifies direct and indirect voting rights.[5] We focus on the year 2015 and, in line with international accounting standards, we limit our definition of control to the case in which a corporate shareholder (i.e., we exclude non-corporate entities) can command the majority (i.e., strictly more than 50%) of the voting rights in another company. As the latter company is majority-owned by the former, we call it a "subsidiary" of the corporate shareholder.[6]

The chosen notion of majority control allows us to identify, for each company in the dataset, the presence (if existing) of a unique corporate "Global Ultimate Owner" (GUO), which is a controlling parent company. The dataset also reports, for each company, information on the "Immediate shareholder" (ISH), that is, the corporate entity that directly controls (at least 50% of) the company under analysis. Starting from the two relations of GUO and ISH, we follow Sonno (2025) and create an algorithm that first identifies the boundaries of a BG as the set of all subsidiaries having the same GUO, and then retrieves the hierarchical distance of a subsidiary from its parent (details on this procedure are described in Appendix A.1).

This procedure generates a dataset of 2,901,466 parent companies controlling 5,676,289 subsidiaries in 2015. For our analysis, we also require additional details on the type of ownership links, which might yield different subsamples of the dataset depending on the information available in the data. Using the Orbis dataset, also produced by Bureau Van Djik, we combined information on the country of incorporation of each company to distinguish international links from domestic ones and map the geographic distribution of the BGs. To analyze the hierarchical layer distribution of subsidiaries, we also require the exact position in the hierarchical layer to be computable. These limitations on data availability slightly reduce the number of subsidiaries to 5,653,433. Moreover, to characterize the BGs, we need information on the industry (at three-digit NAICS revision 2002 of each entity in our database). Orbis does not contain complete information on the industry of activity for all entities, and therefore we lose some 20% of observations in the match with the three-digit NAICS sectors. This is the sample we use except for descriptive statistics, which are based on the

[4] When a single legal entity operates more than one productive plant (multi-plant firm) or is organized internally through multiple divisions, those plants or divisions are sometimes considered by Orbis as different entities and flagged as branches. Since the focus of this work is on the organizational choices of BGs in terms of the control hierarchy, we focus exclusively on the legally independent entities and drop branches from our dataset.

[5] See Appendix A.1 for details on the data. In general, corporate control can be derived by a direct, indirect or consolidated concentration of voting rights (Faccio and Lang 2002; Chapelle and Szafarz 2007; Del Prete and Rungi 2017). For example, accompany H owns 100% of the voting rights (VR) in companies A and B. Company A also owns 70% of the VR in X and 30% of those in Y. Finally, company B owns another 40% of the VR of Y. In this case, H is able to control, directly or indirectly, more than 50% of the voting rights in all the other companies. In fact, H enjoys direct control of A and B, indirect control of X through A, and consolidated control of Y through A and B. This is known as the principle of the Ultimate Controlling Institution in the OECD FATS Statistics (or Ultimate Beneficial Owner in UNCTAD data).

[6] The notion of control stems from OECD (2005), UNCTAD (2009), EUROSTAT (2007). This notion of control neglects cases in which affiliates are *de facto* controlled through minority ownership, as well as cases in which control derives from market advantage (e.g., monopsony) or government regulations (e.g., 'golden share'). This notion of control applies equally to domestic and multinational BGs, and it allows for straightforward comparison with official statistics, as it is commonly used for foreign subsidiaries (Eurostat or OECD FATS) and for international taxation (IAS, IFRS).

Table 1. Sample size.

Number of groups	2,901,466
Number of groups with NAICS	2,250,891
Number of parents with only one sub	2,182,934
Number of subs	5,676,289
Number of subs with layer	5,653,433
Number of subs with NAICS	4,580,899
Number of subs with routinizability	4,550,351

Note: Breakdown of the sample size with details on the number of observations for which we find layer, NAICS sector, and routinizability.

broader sample.[7] A breakdown of the different sample compositions can be found in Table 1.

Some combinations of cross-holdings by shareholders might yield ambiguous positions of subsidiaries across layers. Yet, thanks to the algorithm's efficiency, we can unequivocally assign a layer to approximately 99.6% of the subsidiaries. Nevertheless, to evaluate the robustness of our algorithm in identifying BGs and their hierarchical structure, we also experiment with an alternative approach embedded in network theory. The method has been developed by Rungi et al. (2017) and uses data from the Ownership section of the Orbis Database (a different section with respect to the Historical Ownership Database), thus relying on the original bilateral ownership information sourced by the Bureau Van Dijk from different national sources. For each company recorded, Orbis collects all information available on its ownership structure. We arrange this information in matrix form and then launch the algorithm developed by Rungi et al. (2017), simulating a voting rule in the presence of interlocking assemblies of shareholders. The algorithm detects concentrations of voting rights that allow for strategic coordination, a piece of information that we use to delimit the boundaries of each BG, properly identify the HQs and, thus, assign each subsidiary to its hierarchical position.[8] Reassuringly, both approaches, starting from different initial datasets, and using different methods to identify BGs' boundaries and structure, produce similar results.[9]

Table 2 describes the geographic coverage of the dataset, which spans all countries in the world. We report the number of BGs (equivalent to the number of HQs) and the number of subsidiaries in each country or geographical area. In addition, we specify the number of multinational BGs (i.e., the number of parents that also control subsidiaries abroad) and the number of subsidiaries that are ultimately controlled by a foreign entity or a domestic parent. Most parents and subsidiaries recorded in the data are incorporated in the USA or in Europe. Moreover, unsurprisingly, Europe as a geographical area displays a larger share of multinational groups, given that around half of the subsidiaries of European BGs are located in predominantly European foreign countries.

[7] Going from the broader sample to the restricted sample, we do not induce any self-selection regarding relevant characteristics, in particular, hierarchical differentiation and country of incorporation.

[8] A peculiarity of Rungi et al. (2017) is that it starts from all the ownership paths through which a management decision could run; the approach correctly manages and, thus, take notes of cases of cross-holdings, ownership cycles and consolidation of voting rights across otherwise fragmented webs of equity stakes. In this framework, the hierarchical distance between a parent company and any of its subsidiaries (layer) is thus defined as the shortest ownership path in the network that connects them in an ownership space.

[9] In Appendix A.2, we present the two approaches in more detail and compare them in terms of aggregate descriptive statistics, and we replicate the main evidence on routinizability and a few robustness exercises on the sample of BGs identified through the Rungi et al. (2017) algorithm.

Table 2. Geographic coverage.

	BGs				Subsidiaries			
Region or country	**All**	**%**	**Multinational**	**%**	**All**	**%**	**Foreign**	**%**
Africa	6698	0.24	4852	2.23	41,928	0.74	26,175	2.55
Japan	5357	0.19	4069	1.87	27,844	0.49	4345	0.42
Other Asia	81,855	2.90	23,767	10.92	367,740	6.50	173,615	16.94
Australia	83,677	2.96	3789	1.74	188,346	3.33	28,329	2.76
EU 28	718,057	25.43	115,611	53.10	1,899,408	33.56	506,157	49.38
Other Europe	62,051	2.20	16,258	7.47	154,332	2.73	37,880	3.70
Latin America	31,892	1.13	19,183	8.81	95,995	1.70	65,223	6.36
Russia	53,472	1.89	1051	0.48	159,973	2.83	53,632	5.23
USA	1,743,710	61.74	22,209	10.20	2,628,317	46.44	99,213	9.68
Rest of the world	37,423	1.33	6923	3.18	96,042	1.70	30,510	2.98
Total	2,824,192	100.00	217,712	100.00	5,659,925	100.00	1,025,079	100.00
Country not assigned	77,274				16,364			

Notes: The table details the number of observations of subsidiaries and BGs in a list of regions or countries. It also specifies the number of BGs that control at least one subsidiary abroad (multinational) and subsidiaries that are controlled by foreign parents. For regional aggregation, we have used the online version of the United Nations publication Standard Country or Area Codes for Statistical Use. Under the label "Country not assigned," we classify the observations for which Orbis does not identify a country. Percentages are computed excluding observations located in these countries.

3. THE HIERARCHICAL STRUCTURES OF BGs

This section presents three novel stylized facts about BGs based on our dataset. First, there is a lot of variation across BGs in the number of subsidiaries, the number of layers, and the distribution of subsidiaries across layers. Second, within BGs, subsidiaries engage in increasingly routinizable activities as their hierarchical distance from the parent company increases. Third, within BGs, patenting activity is primarily concentrated at higher hierarchical layers closer to the parent company, and subsidiaries at lower hierarchical layers cite those at higher hierarchical layers much more often than vice versa.

3.1 Hierarchical structures across BGs

In our dataset, there is a lot of variation in the BGs' hierarchical structures in terms of the number of subsidiaries and layers, as well as in the distribution of subsidiaries across layers. Table 3 shows that the vast majority of subsidiaries (68%) are located in the first layers. This feature is relatively more prevalent in domestic subsidiaries (76% of them) rather than in foreign ones (with 42% and 28% of them at the first and second layers respectively). On average, the first four layers contain 97% of all the subsidiaries, although again, multinational groups appear to be distributed across relatively deeper structures, with 91% of them concentrated in the first four layers (97% is reached after six layers).

Table 4 looks at the average number of subsidiaries per layer in further detail by comparing BGs with different numbers of layers. Along the columns, we see that, for any given overall number of layers in which the BG is organized, the number of subsidiaries decreases on average as one moves down the hierarchy, with more subsidiaries closer to the parent company and fewer in lower layers. Along the rows, the average number of subsidiaries in any given layer increases as the BGs' number of layers increases, which is expected as BGs with more layers are on average bigger.

Figure 1 summarizes the distribution of subsidiaries per group. The Panel A shows that around 75% of groups in our data have very simple structures, with one parent controlling only one subsidiary whereas around 20% of the groups have between 2 and 5 subsidiaries. On the other tail of the distribution, around 0.1% of the parents in the sample have more

Table 3. Hierarchical distance.

Layer	Domestic Subsidiaries	%	Foreign Subsidiaries	%	All Subsidiaries	%
1	2,000,151	76.11	357,147	42.39	2,357,298	67.92
2	417,260	15.88	233,214	27.68	650,474	18.74
3	129,417	4.93	121,626	14.44	251,043	7.23
4	46,818	1.78	60,661	7.20	107,479	3.10
5	18,551	0.71	31,742	3.77	50,293	1.45
6	7571	0.29	17,170	2.04	24,741	0.71
7	3878	0.15	9726	1.15	13,604	0.39
8	1867	0.07	4593	0.55	6460	0.19
9	1128	0.04	2563	0.30	3691	0.11
10	582	0.02	1515	0.18	2097	0.06
> 10	735	0.03	2608	0.31	3343	0.10
Total	2,627,958	100.00	842,565	100.00	3,470,523	100.00

Notes: The table shows the number of subsidiaries per layer, distinguishing between domestic and foreign subsidiaries. It excludes BGs with only one subsidiary.

Table 4. Number of affiliates per layer across BGs.

	Maximum layer of the BG										
Avg. subs per layer	1	2	3	4	5	6	7	8	9	10	> 10
1	1.37	3.26	6.24	11.35	20.76	30.97	41.74	45.91	47.57	68.38	81.75
2		2.18	6.13	11.44	19.70	29.01	38.75	43.86	62.43	40.15	63.72
3			2.80	8.36	16.52	23.65	34.91	40.39	45.03	44.08	54.14
4				3.06	9.32	17.17	26.21	30.09	36.17	28.34	43.35
5					3.32	9.24	17.27	21.95	27.94	18.64	37.97
6						3.36	9.32	14.41	19.74	16.09	32.62
7							3.80	8.11	13.58	14.03	31.37
8								3.05	8.57	10.46	19.32
9									3.40	7.54	16.68
10										2.64	13.46
> 10											24.22
N. of BG	2,739,149	126,168	24,172	6811	2571	1157	618	350	207	91	138

Notes: The table shows the average number of subsidiaries per layer for BGs having different maximum layers. If BGs with only one subsidiary were excluded, the mean number of affiliates in the first layer for BGs with only one layer would be 2.81.

than 100 subsidiaries each. The Panel B of Figure 1 shows that, on average, groups organized in a larger number of hierarchical layers (reported on the horizontal axis) tend to have more subsidiaries (see the central bar of the box plot, i.e., the median in the distribution of the number of subsidiaries). Yet, as shown by the outliers of the box plots, there exist relevant exceptions: some groups with only a few layers of control have more than 100 subsidiaries, while some groups with fewer subsidiaries arrange them across several layers.

To capture this high degree of heterogeneity in the distribution of subsidiaries within and across groups through a single summary statistic, we look at the skewness of the distribution of its subsidiaries across layers for BGs with at least two layers (Wellman et al. 2020). When skewness is negative, the distribution is left-skewed and subsidiaries are denser at lower layers (those with larger ℓ), leading to a pyramid-shaped hierarchy. When skewness is

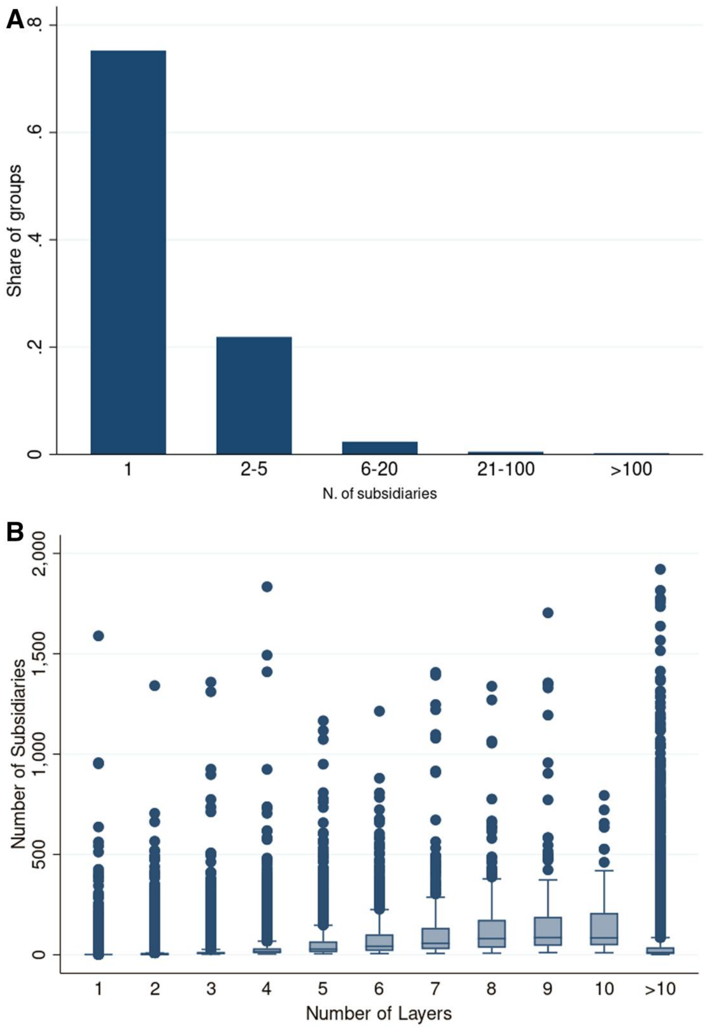

Figure 1. Hierarchical description. Panel A shows the distribution of BG size in terms of number of subsidiaries. Panel B presents 11 box plots of groups' size conditional on the number of layers their hierarchies display. We have excluded the 13 BGs with more than 2000 subsidiaries so as to make the figure more legible.

positive, the distribution is right-skewed and subsidiaries are denser at higher layers (those with smaller ℓ), leading to an inverse pyramid-shaped hierarchy. The closer the skewness is to zero, the more symmetric the distribution is, leading to a diamond-shaped hierarchy.[10]

Figure 2 shows that the distribution of BGs across skewness levels is unimodal, with substantial hierarchical differentiation. In particular, in the full global sample, the skewness of

[10] Recall that the unique assignment of a company to a BG is based on its 'Immediate shareholder' (ISH), which is the corporate entity that directly controls $>50\%$ of the company. This implies that going downwards from the GUO, the building blocks of any hierarchy are either a one-to-one or a one-to-many vertical ownership relationship. Independently of their shapes, all hierarchies are ultimately made of different combinations of those primitive blocks.

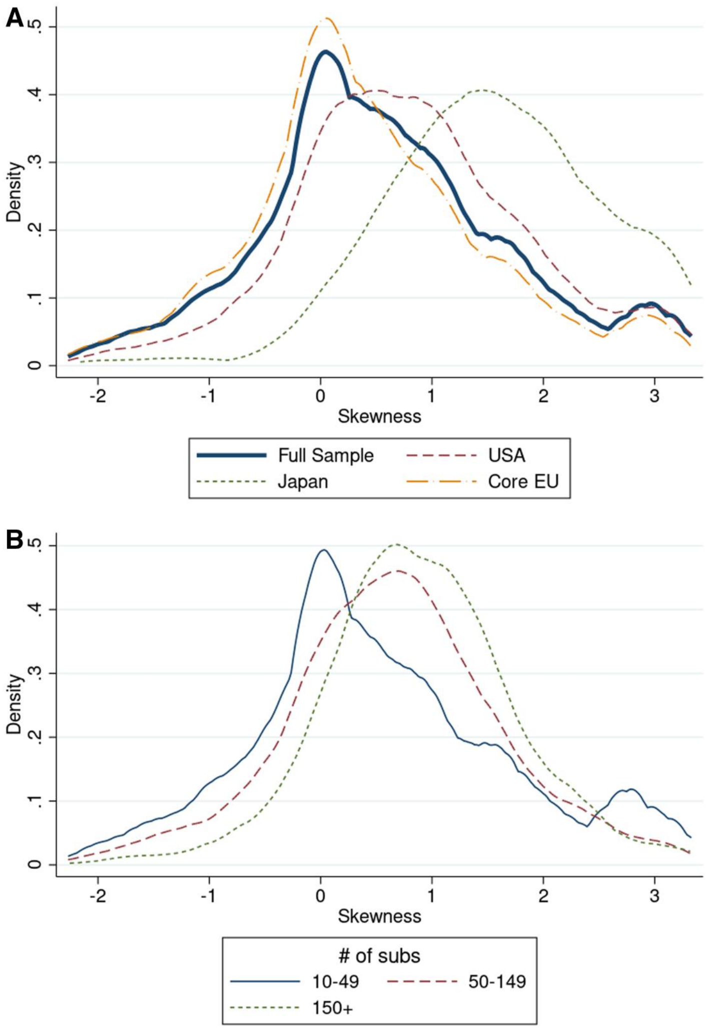

Figure 2. Skewness distribution. The figure shows the skewness distribution of BGs' hierarchies. BGs are grouped by geographical area in Panel A, while in Panel B, they are grouped by size. BGs with less than two layers or 10 subsidiaries are dropped because their skewness can only take a few values and thus alter the densities.

the average distribution of subsidiaries across layers ranges from –2.26 to 3.32 with a mean of 0.19 and a standard deviation of 0.87.

A feature worth noticing is that pyramid-shaped hierarchies (those with negative skewness) are by no means the rule. This is most readily seen by classifying hierarchies with skewness half a standard deviation below zero as "pyramids," those with skewness half a standard deviation above zero as "inverse pyramids," and the rest with skewness within half a standard deviation from zero as "diamonds." The Panel A of Figure 3 shows that, in the full sample, those three classes contain around 20%, 50% and 30% of all BGs, respectively. The Panel B of Figure 3 also highlights that inverse pyramids' dominance is more pronounced in BGs with more subsidiaries. For example, the hierarchies of BGs with 50—149 subsidiaries are around 10% pyramids, 30%

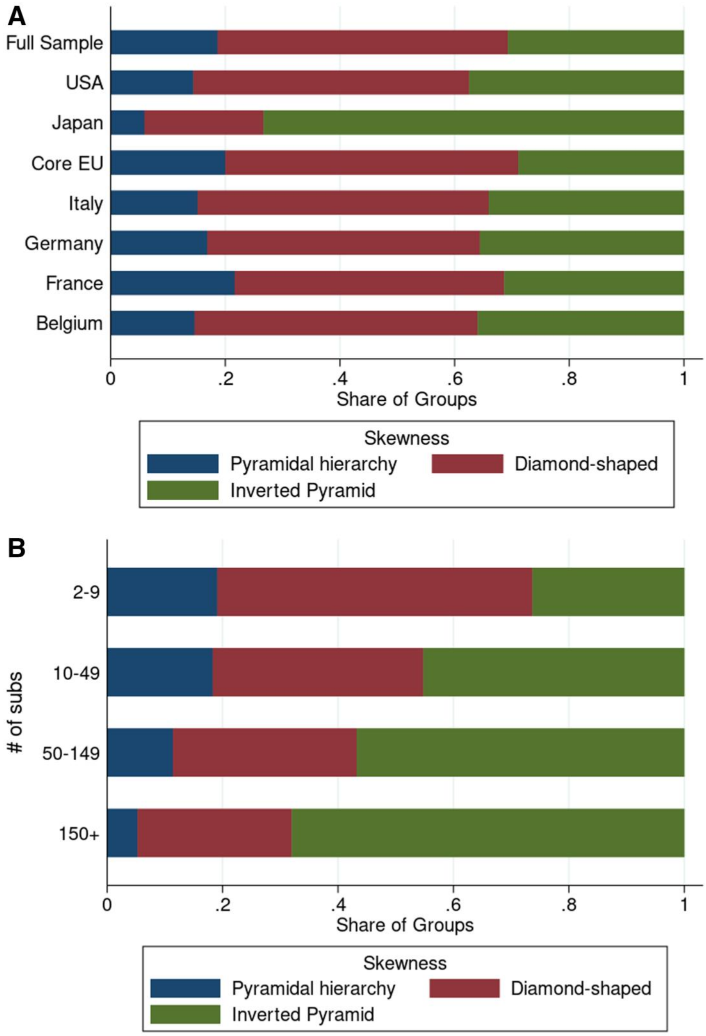

Figure 3. Skewness bars. The figure shows the frequency of the three types of hierarchical structures. In Panel A, BGs are grouped by geographical area, while in Panel B, they are grouped by size. BGs with less than two layers are excluded because their skewness can only take value 0 and thus alter the graphs.

diamonds and 60% inverse pyramids; those with more than 150 subsidiaries are around 5% pyramids, 20% diamonds and 75% inverse pyramids. These patterns are reflected in the mean and the standard deviation of skewness (the Panel B of Figure 2). In BGs with 50–149 subsidiaries mean skewness equals 0.65 with a standard deviation of 0.96, whereas in BGs with more than 150 subsidiaries mean skewness equals 0.88 with a standard deviation of 0.94, compared with 0.19 and 0.87 in the global sample.[11]

[11] Several subsidiaries in the first layer could be financial shells or regional HQs. To check the sensitivity of the skewness patterns reported in Figure 3 to the thickness of the first layer, we have restricted our sample to BGs with at least three layers, and then excluded from our computations the subsidiaries in the first layer. While we observe a decrease in pyramidal shapes

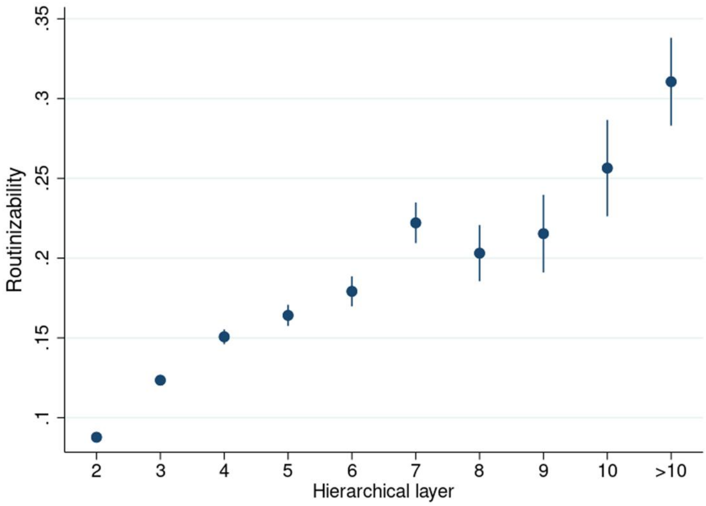

Figure 4. Routinizability over hierarchies. The graph shows the coefficients and 95% confidence intervals obtained from a regression of the routinizability index of each subsidiary on the hierarchical layer of the same subsidiaries, including group and host country FE, and using robust standard errors. Regression results are reported in Table B2. Layer 1 constitutes the omitted category.

Considering the Panel A of Figure 3 again, we see substantial cross-country heterogeneity. Pyramids are less frequent in the USA (18%) and even less in Japan (6%) than in the full global sample. However, diamonds are the dominant shape in the USA (45%) while in Japan (17%) the lion share goes to inverse pyramids (72%). In core European countries, BGs exhibit a more balanced distribution across shapes, roughly in line with the full global sample, though the frequency of pyramids is slightly higher than in the latter. These patterns are reflected in the mean and the standard deviation of skewness. In the USA and the Europe, mean skewness equals 0.19 and 0.14 with standard deviation 0.87 and 0.85, whereas in Japan mean skewness equals 1.15 with a standard deviation of 1.10. Hierarchical differentiation also varies across European countries, with France featuring a larger frequency of pyramids than Germany, Italy and Belgium.[12]

3.2 Hierarchical task complexity within BGs

Figure 4 plots the coefficients with the corresponding 95% confidence intervals estimated regressing for each subsidiary, its index of routinizability on a set of eleven-layer dummies (for $\ell = 1, \ldots, 10$ and $\ell > 10$) plus parent and country fixed effects.[13] The index is constructed by exploiting a specific question Q.25b in the survey questionnaire developed for the Princeton Data Improvement Initiative (PDII) as described in Blinder and Krueger (2009, 2013). This question asks respondents: "How much of your workday involves

and an increase in both inverse pyramids and diamonds, the country and size patterns in Figure 3 are confirmed no matter whether we include or exclude the subsidiaries in the first layer.

[12] One may be concerned that many subsidiaries situated at the periphery of the hierarchical structure without control over other subsidiaries could influence the overall distribution of skewness. This concern is addressed in Appendix B, where some real-world examples are also reported (see Appendix B.1). In particular, Figure B5 in Appendix B.2 replicates Figure 3 after excluding from the computation of skewness the subsidiaries that do not directly control other subsidiaries. The two figures exhibit similar patterns.

[13] Table B2 in Appendix B.3 reports the econometric model and the regression results with robust standard errors behind Figure 4.

carrying out short, repetitive tasks? Would you say ... " There are four possible answers: "Almost all the time—1," "More than half of the time—2," "Less than half of the time—3," "Almost none of the time—4." We compute the mean answer by the three-digit NAICS 2002 sector where respondents are employed and use it as an index of routinizability of the industry where a subsidiary is active. For easier interpretation, we have reversed the original ranking to associate a higher index value with a higher degree of routinizability. We estimate routinizability for 4,550,351 subsidiaries in our sample, obtaining a mean of 2.44, a standard deviation of 0.47, a minimum of 1.33 and a maximum of 4.00.

The figure shows a clear hierarchy of task complexity within BGs as the average routinizability of the tasks performed by a subsidiary in a given layer positively correlates with the distance of its layer from the parent.[14] See also Appendix E, where we show a lack of statistical association between hierarchical layers and contractability. Although the latter is a traditional driver of firms' boundaries, results show that it is unrelated to the organizational structure of BGs.

3.3 Hierarchical knowledge creation within BGs

In addition to the hierarchy of tasks, Figure 5 reveals a specific pattern in patent applications and citations across subsidiaries. Patenting is concentrated either in the HQ or in subsidiaries at layers close to it. This is consistent with the idea of a hierarchy of knowledge creation

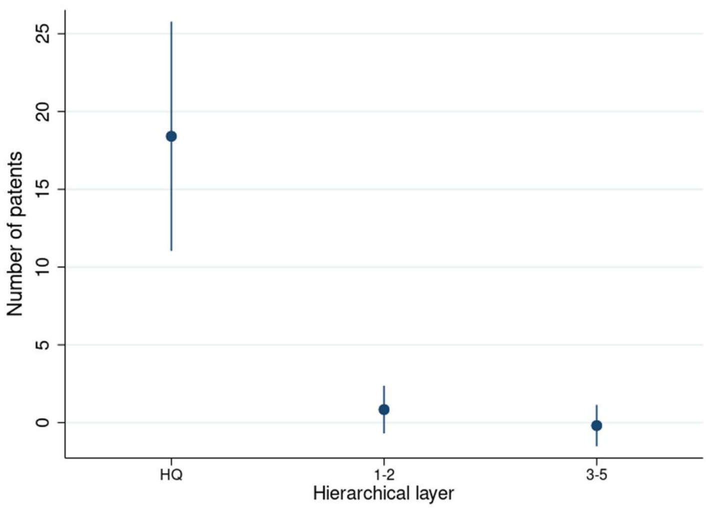

Figure 5. Innovation activity: number of patents. The graph shows the coefficients and 95% confidence intervals after a regression of the numbers of patents of each HQ and subsidiary on the hierarchical layer of the same HQ or subsidiary, including group and host country FE, and using robust standard errors. Layers greater than 5 constitute the omitted category.

[14] A potential concern is that this finding may mask some underlying heterogeneity within BGs. In Appendix B.2, we thoroughly test the robustness of this evidence by conducting various robustness checks that account for affiliates' ownership links and exclude affiliates that do not control other subsidiaries. In addition, in the same Appendix, we replicate this analysis using size controls for both BGs and subsidiaries, and then separating MNEs and domestic BGs. Reassuringly, the results are robust and confirm the pattern shown in Figure 4.

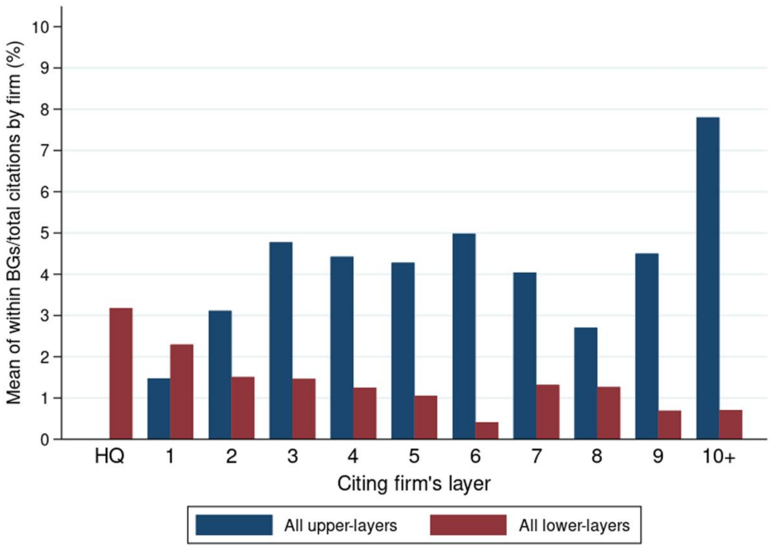

Figure 6. Innovation activity: upper and lower citations. The bar chart shows the average ratio of within-BG citations to total citations received by a firm, by the citing firm's hierarchical layer. Total citations encompass all citations from other firms, both within and outside the firm's BG. This ratio distinguishes within-BG citations of firms in layers (strictly) above the citing firm from those (strictly) below. It should be noted that, for the HQ, the number of citations of affiliates in upper layers is zero by construction.

and diffusion within BGs, as patent citations can be considered as indicators of knowledge flows among partners (Fadeev 2023).[15]

To generate the figure, we have merged information on individual BGs with data on firms' patents and citations retrieved from the Orbis Intellectual Property Database. We have then constructed an original database of firms' networks based on patent-citation links. The database includes around 168,000 firms in 123 countries that have received at least one citation for one patent application, which provides us with 1.4 million patent applications and a network of 9.5 million firm-patent citations.

We have used this dataset to track firms that are part of a BG (as parent or subsidiary) and to identify the layers in which we observe patenting. Finally, we have regressed the number of cited patents attributed to the HQ and its subsidiaries on their hierarchical layer, including group and host country fixed effects and robust standard errors. Layers greater than five constitute the omitted category, given that the presence of cited patents in subsidiaries placed further below along the hierarchy is sparse. Figure 6 shows that, on average, patenting activity is concentrated either in the HQ or in subsidiaries at layers close to it.[16]

[15] Specifically , Fadeev (2023) shows how patent citations primarily originate from business partners, therefore suggesting that they reflect deliberate knowledge sharing rather than unintended spillovers.

[16] We have also explored the "quality" of patents as proxied by citations. Appendix C.1 shows that high-quality patents (i.e., those that are cited not only by other patents belonging to the same firm) are located in the HQ or in layers close to it, while patents placed at layers further down the hierarchy, if anything, mostly cite themselves. Specifically, Figure C1 shows that the average ratio of self-citations out of citations per layer rises as one moves down the hierarchy. Moreover, in Appendix C.2, we complement these results with evidence of interactions among firms across the BG's hierarchy: we document a positive and significant correlation (in OLS with fixed effects) between productivity shocks in bottom layers and the change of productivity in the HQ. This is consistent with the idea that hierarchical choices are not exogenous to the overall performance of a BG.

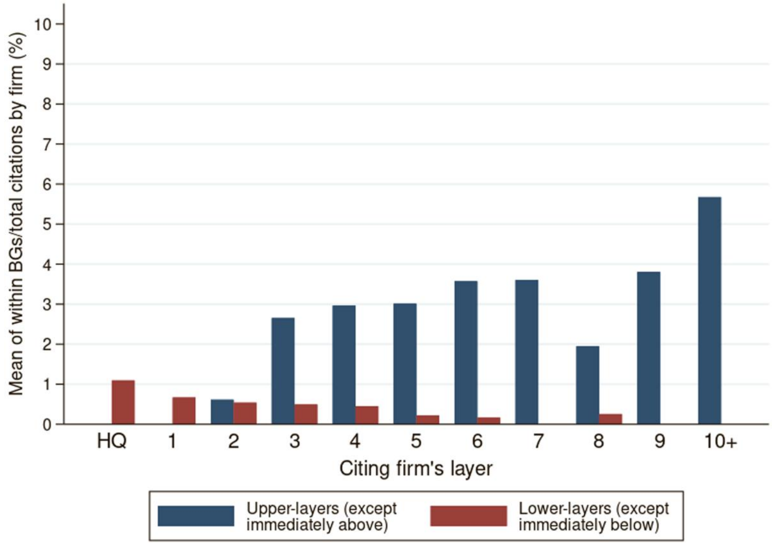

Figure 7. Innovation activity: upper-and lower-citations excluding adjacent layers. The bar chart shows the average ratio of within-BG citations to total citations received by a firm, by the citing firm's hierarchical layer. Total citations encompass all citations received from other firms, both within and outside the firm's BG. This ratio distinguishes within-BG citations of firms in layers (strictly) above the citing firm from those (strictly) below. It should be noted that, for the HQ and the first layer, the number of citations of affiliates in the upper layers is zero by construction. Citations of firms in the layer immediately above and immediately below are excluded.

We have further exploited the unique features of our original patent dataset to study the hierarchy of knowledge diffusion within BGs by looking at patents' citations between subsidiaries at different layers. Figure 6 reports the average ratio across BGs of within-BG citations to total BG citations for patents registered at all layers above or below any given layer. It shows that lower layers are more likely to cite higher layers than vice versa. Figure 7 looks at the same ratio after excluding citations between adjacent layers. It shows that lower layers are still more likely to cite higher layers than vice versa. Hence, citations are not necessarily associated with direct ownership links.

3.4 Discussion and related literature

Our evidence documents a large variation across BGs' hierarchical structures in terms of the number of subsidiaries, the number of layers, and the distribution of subsidiaries across layers. The literature has highlighted how ownership links can be designed to create a network through which resources flow between the subsidiaries of a BG. This "internal market" allows the HQ to circumvent the imperfections of the external market and promote a more efficient resource reallocation in a changing business environment. Three dimensions of internal resource reallocation have received particular attention. The first concerns capital (Khanna and Palepu 2000; Khanna and Rivkin 2001; Almeida and Wolfenzon 2006; Almeida et al. 2011; Belenzon and Schankerman 2013). Financial resources can be moved across subsidiaries to fund those with higher returns on investment or bail out those in financial distress. The second dimension regards labor (Huneeus et al. 2021; Cestone et al. 2023). Employees can be reallocated across subsidiaries to respond to demand shocks and changes in business opportunities. The third dimension concerns taxable income (Lewellen and Robinson 2013), which can be shifted across subsidiaries for tax arbitrage purposes.

Our analysis highlights a fourth important dimension of internal resource allocation in a BG, in which higher-layer subsidiaries perform more complex tasks and register more cited patents than lower-layer subsidiaries, thereby fostering an efficient allocation of knowledge creation, diffusion and application across subsidiaries. This is consistent with the seminal work of Bartlett and Ghoshal (1989), who emphasize the importance of allocating knowledge by managing the increasingly complex flow of information within international organizations through various processes, such as the rotation of personnel, the setup of informal communication channels, and the use of task forces and committees. The latter contribution started a literature in which knowledge flows from parent companies to their subsidiaries (Gupta and Govindarajan 1991; Monteiro et al. 2008) and vice versa from subsidiaries to their parents (Bresman et al. 1999; Nair et al. 2015) have been documented. Recent works have also highlighted the possibility of transferring knowledge within BGs in directions other than the direct ownership links. Specifically, the transmission of knowledge between peer subsidiaries ("lateral knowledge transfer") has increasingly gained attention (Ciabuschi et al. 2011; Li and Lee 2015; Crespo et al. 2020; Garg et al. 2022; Wu et al. 2022). This lateral direction of knowledge flow seems to be the result of various practices adopted by BGs, such as international temporary task forces and inter-unit committees, training and workshops, personnel exchanges, and the adoption of corporate-wide intranet systems (Subramaniam and Venkatraman 2001; Björkman et al. 2004; Persson 2006; Ciabuschi et al. 2011; Crespo et al. 2020).

We add to these channels patents cross-citations, with respect to which a clear pattern emerges from the data: knowledge is predominantly created in the HQ (or upper-layer subsidiaries) engaging in more complex tasks and subsequently disseminated down the entire hierarchy.

In what follows, we build on the foregoing strands of literature and our empirical findings to rationalize a BG's hierarchical structure as a means to promote the efficient creation, diffusion and application of knowledge through the optimal design of an internal knowledge market. In particular, adopting the idea put forth by Garicano and Rossi-Hansberg (2015) for hierarchies of employees within companies, we conceptualize knowledge as the ability to solve the problems that naturally arise in any production process.

4. A KNOWLEDGE-BASED ORGANIZATIONAL THEORY OF BGs

The main idea is that, as in Garicano and Rossi-Hansberg (2015), knowledge resides with the top management and is communicated to underlings by advising their problem solving; in a BG, knowledge emanates from the HQ and is communicated to subsidiaries by advising their own problem solving. Advising time for the HQ comes, however, in limited amount. Moreover, some problems are harder to solve than others. Their solution has higher upside potential in terms of the BG's performance, but it also absorbs more advising time. The BG's optimal hierarchy then arises from the HQ's profit-maximizing allocation of advising time across subsidiaries solving problems of different difficulty. Together with the nature of the problems, three primitive parameters play a key role in determining the optimal hierarchy. These capture the HQ's efficiency in producing, in advising knowledge creation, and in handling the associated communication across subsidiaries. Crucially, thanks to the internal knowledge market, the communication of knowledge within the BG is not restricted to following direct ownership links.

4.1 Knowledge creation

Consider a BG consisting of a parent and an endogenous number of subsidiaries. The parent owns the "blueprints" of a large portfolio of final products and has to decide how many of them to produce, and how to organize their production through subsidiaries. Each product faces an isoelastic demand $y = Ap^{-\sigma}$, where y is output, p is price, $\sigma > 1$ is the demand elasticity, and $A > 0$ is a demand shifter. A product's revenues are then given by $R(y) = py = A^{1/\sigma}y^{(\sigma-1)/\sigma}$. Both σ and A capture market characteristics that are common to all products, and thus to all BGs.[17]

The parent has exclusive knowledge of the production possibilities of each blueprint but, in order to turn them into actual production, a product-specific problem has to be solved. The problem comes in different versions, indexed by $\wp$ and ranked $\{1, 2, \ldots\}$ in decreasing order of difficulty $D(\wp) = e^{-\theta\wp}$ with $\theta > 0$. This is the same for all products and is not BG-specific. The parent knows how to solve any version of any problem, but it is not directly involved in production, which it delegates to the subsidiaries by deciding which blueprints, which problems and which versions to assign them. Direct or indirect majority ownership of the subsidiaries is needed for the parent to be in control of their use of its blueprints.

A subsidiary consists of a problem solver ("executive") and a team of producers ("employees") whose number increases with the level of output. If the executive does not solve the problem decided by the parent, the employees in her subsidiary cannot produce, and their productivity is at zero. If she solves the problem, their productivity depends on the assigned version's difficulty, with the former increasing with the latter. Specifically, if the executive solves version $\wp$, her employees' productivity is $\Omega(\wp) = \omega D(\wp) = \omega e^{-\theta\wp}$, where $\omega > 0$ is a parent-specific component capturing the fact that some BGs have better production possibilities than others.

All employees have the same skills. Their wage is normalized to 1, and at this wage, their supply is infinitely elastic. In contrast, executives come in different ability types. Solving more difficult versions requires higher ability that not all executives have. Executive types are indexed by m and ranked $\{1, 2, \ldots\}$ in decreasing order of ability $M(m) = e^{-m}$, which is defined as the capacity to solve problem versions $\wp = 1, \ldots, m$ and thus implies that a problem version $\wp$ can be solved only by an executive of type $m = 1, \ldots, \wp$. Executives have only a limited amount of time they can devote to problem-solving. In this amount of time each of them can solve at most one problem. Executives of higher ability are more expensive with $W(m) = wD(m) = we^{-\theta m}$ denoting the fixed cost of hiring an executive m whose problem-solving capacity allows employees in her subsidiary to achieve productivity $\Omega(m)$. At hiring cost $W(m)$ the supply of executives m is also infinitely elastic. Executive remuneration per efficiency unit w is a market characteristic common to all BGs.

On top of adequate ability, in order to solve her subsidiary's problem version, the executive also needs advice from the parent, which can transmit its problem-solving knowledge either directly or indirectly through executives in other subsidiaries. In the latter case, advice can be obtained from any of the BG's subsidiaries no matter whether or not the advised subsidiary is owned (directly or indirectly) by the advising one. However, the parent's knowledge can be transmitted to an executive only through executives of higher ability.

Advising is time consuming for the advisor and the amount of time needed increases with the difficulty of the problem version to be solved and decreases with the advisee's ability. Specifically, in order to solve a problem version of difficulty $D(\wp)$, an executive of ability $m \leq \wp$ requires $T(\wp, m) = D(\wp)M(m)^{-\varphi} = e^{-\theta\wp}e^{\varphi m}$ units of advising time. The rate $\varphi > 0$,

[17] See Appendix D for a standard microfoundation of this demand system under monopolistic competition.

at which communication becomes less efficient as the advisee's ability falls, is a parent-specific characteristic capturing different communication efficiency across BGs.

The amount of available advising time is $\widehat{T} = e^{\tau}$ for both the parent and each executive, where $\tau > 0$ is a BG-specific measure of advising efficiency. Offering advice is the only activity of the parent, hence $\widehat{T}$ is its total endowment of time. Differently, for an executive $\widehat{T}$ is an extra amount of time in addition to the one she has for problem solving. For simplicity, the executive's advising and problem-solving amounts of time are not substitutable, and the hiring cost remunerates both.

4.2 Optimal knowledge hierarchy

As in the data, also in the model, the unique assignment of a subsidiary to a parent is based on the corporate entity holding direct majority control.[18] This implies that, going downwards from the parent, the building blocks of any BG's hierarchy are either one-to-one or one-to-many vertical ownership relations. All hierarchical shapes are ultimately made of different combinations of such building blocks. We now characterize the conditions under which the combination of those blocks into the shapes we actually observe in the data can be explained through the lens of our model in terms of efficient knowledge management. In particular, we are interested in understanding how BGs' organizational structures vary as a function of the three parameters that capture their efficiency in producing (ω), advising (τ), and communicating (φ).

A BG's knowledge hierarchy is determined by the way in which the executives of different abilities leading its subsidiaries are arranged in hierarchical layers of problem solvers who are both advisors and advisees. For example, with layers indexed $\ell \geq 0$, in a "pyramid" structure the parent at layer $\ell = 0$ advises directly a smaller number of subsidiaries at top layer $\ell = 1$ and indirectly a rising number of subsidiaries at increasingly lower layers $\ell > 1$; in an "inverse pyramid" structure the parent at layer $\ell = 0$ advises directly a larger number of subsidiaries at top layer $\ell = 1$ and indirectly a diminishing number of subsidiaries at increasingly lower layers $\ell > 1$; in a "diamond" structure there is a threshold layer $\ell > 1$, above which the knowledge hierarchy is a "pyramid" and below which it is an "inverse pyramid."

Creating a layer bears an administrative cost $wF > 0$, which is a market characteristic common to all BGs. The optimal hierarchy can be derived by analyzing the parent's decision to activate the first layer $\ell = 1$ and then iterating the analysis to the next layers. Within a layer, the parent anticipates the profit a subsidiary placed there will earn as the result of an optimal selection among problem versions and executive types hired to solve them. Choices maximize the BG's profit and are made sequentially. Initially, the parent decides whether to create the layer, how many subsidiaries n to activate (i.e., how many product blueprints to implement) and which version $\wp$ of the corresponding problems to solve. Then, it selects the executive type m tasked with solving the subsidiaries' problems. Finally, after the executives have solved the problems, the subsidiaries decide how much output y to produce and earn the corresponding profits from sales.

The main results, derived in Appendix D, are that as long as the demand elasticity is large enough ($\sigma > 2$), the parent assigns the problem versions $\wp$ to layers in decreasing order of difficulty from the top and tasks with solving them the least able executives who can succeed (i.e., those with $m = \wp$) under the advice of the least able executives who can help them (i.e., those with $m - 1$). Moreover, as long as the parent-specific productivity component ω is large enough, the first layer is assigned the most difficult problem versions so that the layer and problem version indexes coincide ($\ell = \wp$). Large enough demand elasticity ensures that, in terms of profit, the productivity gains from solving harder problem versions more

[18] See Footnote 10, where this point is made regarding the data.

than compensate the additional hiring costs for the better executives needed to solve them.[19] High enough parent productivity ensures that when the first layer is assigned and solves the hardest problem version ($\wp$), its profit is not negative.[20] Intuitively, these results are explained by the fact that it is more profitable for the BG that the parent and the executives devote their limited advising time to the most productive problem versions they can help solve.

To summarize, if demand elasticity and parent productivity are large enough, the layer coincides with the problem version, which in turn is matched by the executive type ($\ell = \wp = m$). The feasible number of subsidiaries at generic layer ℓ is then determined by the time constraint for layer $\ell - 1$, which equalizes the advising time needed by the former layer and the latter layer's endowment: $n_\ell e^{(\varphi-\theta)\ell} = n_{\ell-1} e^{\tau}$. Applying the constraint recursively from the parent (for which $n_0 = 1$) yields the number of subsidiaries

$$n_\ell = e^{\ell\tau + \frac{1}{2}\ell(\ell+1)(\theta-\varphi)} \tag{1}$$

for the generic layer ℓ. However, the layer is activated only if it generates positive profit

$$n_\ell \Pi_\ell - wF = n_\ell\left[a\omega^{\sigma-1} e^{-\theta(\sigma-2)\ell} - w\right] e^{-\theta\ell} - wF > 0 \tag{2}$$

where

$$\Pi_\ell = \left[a\omega^{\sigma-1} e^{-\theta(\sigma-2)\ell} - w\right] e^{-\theta\ell}$$

is profit per subsidiary and $a = (A/\sigma)[(\sigma-1)/\sigma]^{\sigma-1}$ is a bundle of demand parameters. Hence, the hierarchy stops at the lowest layer, that makes non-negative profit. This is the cutoff layer ℓ^* such that

$$n_\ell \Pi_\ell - wF \geq 0 \text{ and } n_{\ell+1}\Pi_{\ell+1} - wF < 0 \tag{3}$$

simultaneously hold for $\ell = \ell^*$.

4.3 Pyramids and diamonds

The model has rich implications for a BG's hierarchical structure.[21] The cutoff layer defines the depth of the hierarchy with larger ℓ^* describing a deeper hierarchy with a larger number of layers. As a layer's profit $n_\ell \Pi_\ell - wF$ is an increasing function of ω, a BG's hierarchical depth increases with its production efficiency. Depth also depends on the rate at which the layer's profit falls when its hierarchical distance from the parent rises, which in turn depends on how Π_ℓ and n_ℓ evolve when ℓ grows.

Since higher ℓ is associated with increasingly simpler problem versions and thus increasingly smaller productivity gains from solving them, profit per subsidiary Π_ℓ falls when ℓ rises, with speed determined by the rate θ, at which problem versions become easier to solve, and

[19] A value $\sigma > 2$ finds empirical support in the literature (see, e.g., Head and Mayer 2014).

[20] The exact condition on parent productivity, derived in Appendix D, is $\omega > \left[(w/a)(1 + Fe^{\varphi-\tau})e^{\theta(\sigma-2)}\right]^{\frac{1}{\sigma-1}}$ with $a = (A/\sigma)[(\sigma-1)/\sigma]^{\sigma-1}$. When this condition is violated, the first layer would make losses if it were assigned the most difficult problem version $\wp = 1$. It would then be assigned the next most difficult problem version that made it at least break even. Using $\wp_{\max}$ to denote this problem version, the layer index ℓ would still start from 1, whereas the problem version and executive type indexes would start from $\wp_{\max}$. This situation would only complicate the notation without providing any concrete additional insight.

[21] See Appendix D for a formal derivation.

the rate σ, at which decreasing productivity gains translate into decreasing sales gains. However, as both θ and σ are market characteristics common to all BGs, heterogeneous depth across them cannot be explained by these parameters. It is explained by the different behavior of n_ℓ, as determined by the problem-solving efficiency parameters driving it.

To see this, it is useful to rewrite the advising time constraint $n_\ell e^{(\varphi-\theta)\ell} = n_{\ell-1}e^{\tau}$ in terms of "span of control" defined as the number of advisees per advisor $n_\ell / n_{\ell-1} = e^{\tau-(\varphi-\theta)\ell}$. For $\varphi < \theta$, as ℓ rises, problems become easier to solve more rapidly than their solutions become harder to communicate. As a result, the span of control increases and a given number of upstream advisors can successfully advise a larger number of downstream advisees, which resembles the typical situation emphasized in models of knowledge hierarchies within companies (Garicano 2000; Garicano and Wu 2012; Garicano and Rossi-Hansberg 2015).

In contrast, for $\varphi > \theta$, as ℓ rises, problems become easier to solve less rapidly than their solutions become harder to communicate. In this second case, an advisor's available time gets increasingly crowded so that her span of controls falls as ℓ rises. As a result, a given number of upstream advisors can successfully advise a larger number of downstream advisees in higher layers ($\ell < \tau/(\varphi-\theta)$), while they can advise only a smaller number in lower layers ($\ell > \tau/(\varphi-\theta)$). The number of subsidiaries per layer n_ℓ is then a hump-shaped function of ℓ, as implied by (1). An inverse pyramid arises when the first layer precedes the hump's top ($\tau/(\varphi-\theta) < 1$), a diamond when it follows the hump's top but precedes the cutoff layer ($1 < \tau/(\varphi-\theta) < \ell^*$), and a pyramid when it follows the cutoff layer ($\tau/(\varphi-\theta) > \ell^*$).

Overall, given the same problem-solving challenges (same θ), pyramids are chosen by BGs with high efficiency in advising (large τ) and communication (small φ), inverse pyramids by BGs with low efficiency in advising (small τ) and communication (large φ), and diamonds by BGs of intermediate efficiency. Intuitively, for less efficient BGs in advising and communicating, it is better to keep problem solving, and thus production, on layers closer to the parent.

4.4 Quantitative analysis

To check its empirical relevance, the model can be calibrated to target some key moments of the data and then simulated to see whether its predictions are consistent with other untargeted moments reported in Section 2. Specifically, we exploit the model's equilibrium equations to specify econometric regressions that allow us to obtain structural estimates for the parameters regulating problem difficulty (θ), advising efficiency (τ), and communication efficiency (φ). Then, we use those estimates, together with the model's structure, to calibrate the parameter regulating production efficiency (ω). We choose the BGs' number of layers as the target moments of the data and rely on normalizations and existing estimates borrowed from the literature for the remaining parameters. Finally, we show that the model correctly predicts that BGs organized in a larger number of hierarchical layers tend to have more subsidiaries, and that the dominance of inverse pyramids is more pronounced in BGs with more subsidiaries. We also show that it correctly predicts that left-skewed hierarchies are less frequent in the USA and in Japan than in Europe.

In the case of θ, we recall its definition in $D(\wp) = e^{-\theta\wp}$ as the rate at which the solutions of more difficult problem versions become harder to find. Given $\wp = \ell$, since more difficult problem versions are allocated to higher layers, the model implies that θ is also the rate at which problem version difficulty falls as the layer index ℓ increases. Based on this property, we proxy the difficulty $D(\ell)$ of problem versions addressed at layer ℓ by the inverse of the routinizability of the layer's activities, measured by the same variable as in Figure 4 and

Table 5. Reduced form estimates of θ.

Dependent variable	(1)	(2) log(Routinizability)	(3)
Estimation method	**OLS**	**OLS**	**OLS**
Layer	0.0207***	0.0252***	0.0156***
	(0.000846)	(0.000609)	(0.000784)
Parent FE	Yes	Yes	Yes
Subcountry FE	Yes	Yes	Yes
Cluster	BG	BG	BG
Sample	Excl. 1to1	Excl. 1to1	Excl. 1to1
		Layers 1-6	# Subs ≥ 10
Observations	2,687,600	2,666,804	1,121,548
R^2	0.536	0.541	0.360

Notes: OLS estimations. *log(Routinizability)* is the logarithm of the routinizability of the layer, as from the PDII at the three-digit NAICS 2002 sector level; *Layer* is the degree of separation between the HQ and the corresponding subsidiary. The constant is omitted from the table. All specifications include parent company FE and subsidiary's country FE, and are estimated on the full world sample, excluding BGs with only one subsidiary. Specification (2) only includes the subsidiaries in the first six layers, specification (3) only includes subsidiaries belonging to BGs with ten or more subsidiaries. Standard errors clustered at the HQ level are in parentheses.
*** $p<.01$.

retrieved from Blinder and Krueger (2009, 2013). Given that restating $D(\ell)=e^{-\theta\ell}$ in logs yields

$$\ln(D(\ell)^{-1})=\theta\ell, \tag{4}$$

we estimate θ by regressing the log routinizability of a layer on the layer's index controlling for BG fixed effects and subsidiary host country fixed effects, and clustering standard errors at the HQ level. The estimated θ thus corresponds to the average semi-elasticity of routinizability to the layer index within groups. The estimation results are reported in Table 5 for the full world sample. The table shows that, as one moves down the hierarchy by one layer, routinizability increases by around 2% on average.

As for τ and φ, we estimate them starting from equation (1) as follows. Taking logs and bundling the primitive parameters as $\alpha_1=\tau+(\theta-\varphi)/2$ and $\alpha_2=(\theta-\varphi)/2$ gives

$$\ln(n_\ell)=\alpha_1\ell+\alpha_2\ell^2, \tag{5}$$

which implies that α_1 and α_2 can be estimated by regressing the log number of subsidiaries in a layer on the layer's index.[22] Then, using the estimated α_1 and α_2 together with the previously estimated θ, the corresponding values of τ and φ can be backed out by inverting the definitions of α_1 and α_2 as $\tau=\alpha_1-\alpha_2$ and $\varphi=\theta-2\alpha_2$. The estimates of α_1 and α_2 from regression (5) are displayed in Table 6 for the four groups of countries. Both coefficients are consistently and significantly estimated across the four samples. Sensitivity checks (columns 5 to 8) performed excluding BGs with more than 6 layers show that the estimates are not overly sensitive to extreme distributions in the shape of the BGs.

To estimate τ and φ together with their standard errors, we bootstrap our estimates of α_1 and α_2 on 1000 random subsamples, each including 5% of the relevant BGs. We then compute τ and φ using $\theta=0.0207$ from Table 5, Column (1).

[22] As BGs with only one affiliate are uninformative for this regression, we exclude them from the estimation.

Table 6. Reduced form estimates of α_1 and α_2.

Dependent variable	(1)	(2) log($\mathbf{n}_\wp$)	(3)	(4)	(5)	(6) log($\mathbf{n}_\wp$)	(7)	(8)
Estimation method	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**
Country/area	World	USA	Japan	C. Europe	World	USA	Japan	C. Europe
Layer	0.666***	0.848***	1.313***	0.625***	0.833***	0.968***	1.591***	0.732***
	(0.00851)	(0.0103)	(0.0492)	(0.0147)	(0.00159)	(0.00226)	(0.0271)	(0.00414)
Layer2	−0.0650***	−0.0814***	−0.185***	−0.0721***	−0.141***	−0.154***	−0.298***	−0.121***
	(0.00333)	(0.00525)	(0.0200)	(0.00638)	(0.000923)	(0.00196)	(0.0123)	(0.00232)
Cluster	BG	BG	BG	BG	BG	BG	BG	BG
Sample	Excl. 1to1	Excl. 1to1	Excl. 1to1	Excl. 1to1	Excl. 1to1, layers 1-6	Excl. 1to1, layers 1-6	Excl. 1to1, layers 1-6	Excl. 1to1, layers 1-6
Observations	939,062	384,629	4,768	191,700	935,812	383,794	4,738	191,330
R^2	0.512	0.708	0.535	0.459	0.541	0.729	0.571	0.473

Notes: OLS estimations. $log(n_\wp)$ is the log number of subsidiaries at layer $\wp$; *Layer* is the degree of separation between the HQ and the corresponding subsidiary. Together with the estimated θ from Table 5, the estimate coefficients of *Layer* and *Layer*^{2} allow for the identification of the values of τ and φ. Model estimated for four groups of countries (world, USA, Japan, core Europe) always excluding BGs with only one subsidiary. Specifications (5)–(8) restrict the sample to only include subsidiaries in the first six layers. Standard errors clustered at the HQ level are in parentheses.
*** $p < .01$, ** $p < .05$, * $p < .1$.

Table 7. Bootstrap for τ and φ.

	τ			φ		
Country/area	**5th p.**	**50th p.**	**95th p.**	**5th p.**	**50th p.**	**95th p.**
World	0.668	0.751	0.820	0.114	0.161	0.204
USA	0.858	0.961	1.042	0.137	0.203	0.267
Japan	1.307	1.802	2.529	0.271	0.569	1.088
Core Europe	0.603	0.742	0.831	0.108	0.191	0.251

Notes: Estimates are conducted on 1000 bootstrapped samples covering 5% of the BGs in each country/area. φ is computed using $\theta = 0.0207$ which is our preferred estimate in Table 5, Column (1).

The bootstrapped results are reported in Table 7 for the four country groups, taking the median across the subsamples as point estimate, and the 5^{th} and 95^{th} percentiles of the distribution as confidence intervals. The estimated parameters indicate that US BGs are more efficient than European BGs in advising knowledge creation; while they are similarly efficient in handling associated communication across subsidiaries. Moreover, while Japanese BGs are more efficient than European and US BGs in advising knowledge creation, they are less efficient in handling associated communication across subsidiaries.

To validate the country rankings obtained from the model's estimates, we compare them with those that can be retrieved from related measures in the World Management Survey (WMS). To proxy advising efficiency we take the mean of "Process Documentation," "Performance Tracking," "Performance Review," and "Consequence Management" scores. To proxy communication efficiency we take the mean of "Performance Dialogue" and "Clarity of Goals and Measurement" score. For each category the score ranges between 1 (weak) and 5 (strong). Table 8 shows that the country rankings based on the WMS scores are aligned with the estimated ones reported in Table 7.

Turning to production efficiency ω, we calibrate it through the structure of the model. Specifically, the model loads cross-country variation on the heterogeneity of φ and τ, and within-country variation across BGs on the heterogeneity of ω, while assuming that all other parameters (A, F, w, θ, σ) are the same for all BGs. For calibration, we normalize $A = F = w = 1$, lift the estimated θ from Table 5 and set $\sigma = 5$ following Head and Mayer (2014). We then constrain φ and τ to be the same for all BGs in each country group. For each BG, we retrieve ω follows. We start by considering the cutoff layer condition (3) for the targeted BG, under the chosen values of the other parameters. We then iterate on alternative values of ω until condition (3) delivers the BG's observed cutoff layer, that is, until the maximum layer generating non-zero profit in the model coincides with the actual one.

To show that the model correctly predicts that BGs organized in a larger number of hierarchical layers tend to have more subsidiaries, and that the dominance of inverse pyramids is more pronounced in BGs with more subsidiaries, we focus on US BGs and partition them in percentiles of their cutoff layers' distribution. Moreover, in order to have enough variation across percentiles, we consider only the BGs with at least 10 subsidiaries and restrict the calibration to the top 90, 95 and 99 percentiles. For the US sample, the estimated advising and communication efficiency parameters are $\tau = 0.961$ and $\varphi = 0.203$ (see Table 7). For the targeted percentiles, the observed cutoff layers are 5, 6 and 9, with implied values of ω equal to 2.090, 2.131, and 2.480, respectively. Hence, a higher percentile is associated with higher production efficiency.

According to the model this finding has two implications. The first concerns the number of subsidiaries. This is predicted to be 27, 34, and 44 for the top 90, 95, and 99 percentiles,

Table 8. Comparison with WMS data.

Country	Supervision (τ)	Communication (ϕ)
Japan	0.689	0.648
USA	0.675	0.615
Core Europe	0.638	0.557

Notes: The table shows some proxies of country-level efficiency in knowledge supervision and communication, based on the World Management Survey data. Supervision efficiency is proxied by the mean of "Process Documentation," "Performance Tracking," "Performance Review, "and "Consequence Management" abilities. Communication efficiency is proxied by the mean of "Performance Dialogue" and "Clarity of Goals and Measurement" abilities. Each ability is reported in the WMS survey with an increasing score from 1 (weak) to 5 (strong). Mean scores τ_{WMS} and φ_{WMS} are weighted by country, sector, and firm's size class.

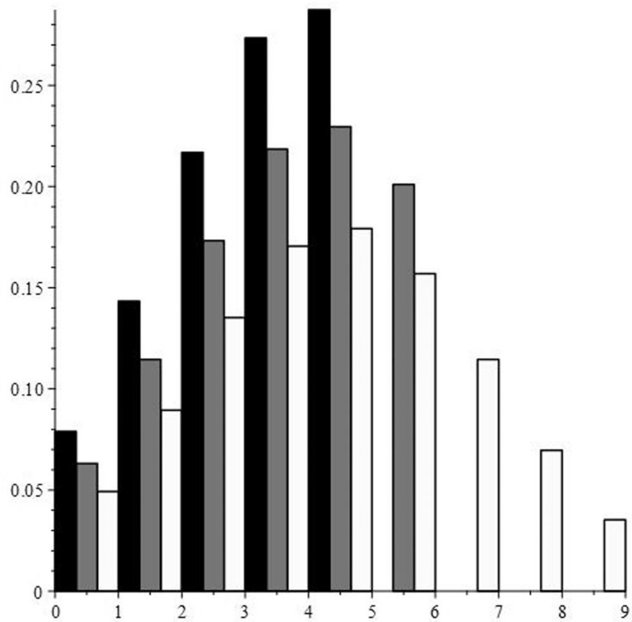

Figure 8. Hierarchies with more layers are less left-skewed. Model simulation. Density of subsidiaries across layers for BGs in the top 90 (black), 95 (gray), and 99 (white) percentiles of the distribution of the number of layers. US BGs with at least 10 subsidiaries. Parameter values: $A = F = w = 1$, $\theta = 0.0207$, $\sigma = 5$, $\tau = 0.961$, $\varphi = 0.203$, $\omega = 2.090$ (black), 2.131 (gray), and 2.480 (white).

respectively, which implies that it increases with a BG's productivity. The second implication refers to the allocation of subsidiaries across layers. This is depicted in Figure 8, which reports the density of subsidiaries across layers, with black, gray and white bars corresponding to BGs in the top 90, 95 and 99 percentiles, respectively. Higher production efficiency makes the hierarchical structure less left-skewed. Considering the two implications together, the model correctly predicts that BGs organized in a larger number of hierarchical layers tend to have more subsidiaries and are less likely to be organized as pyramids.

The foregoing result holds when τ and φ are assumed to be the same for all BGs in a country. One may argue that this is a strong assumption, especially for multinational and

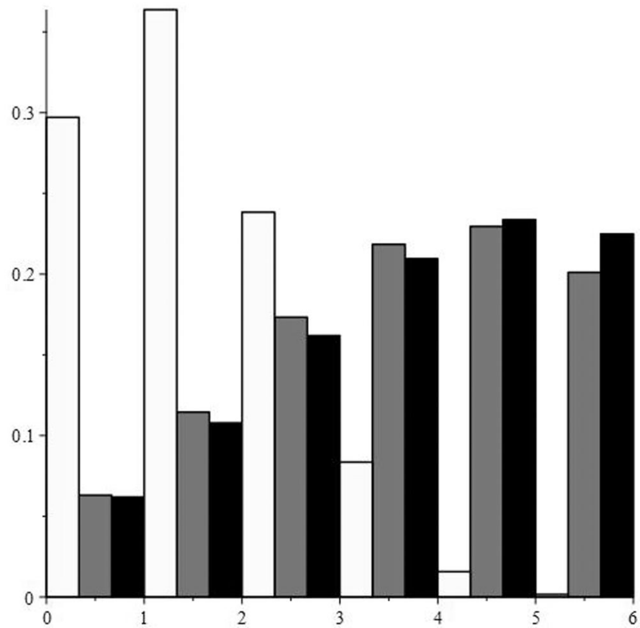

Figure 9. Hierarchies are less left-skewed in domestic than in multinational BGs. Model simulation. Density of subsidiaries across layers for BGs in the top 90 percentile of the number of layers distribution. US BGs with at least 10 subsidiaries. Parameter values: $A = F = w = 1$, $\theta = 0.0207$, $\sigma = 5$; $\tau = 0.851$, $\phi = 0.169$, $\omega = 2.131$ (multinational; black); $\tau = 0.961$, $\varphi = 0.203$, $\omega = 2.131$ (pooled; gray); $\tau = 1.453$, $\varphi = 0.646$, $\omega = 10.655$ (domestic; white).

domestic BGs. To investigate the implications of relaxing this assumption, we rerun regression (5) on the US sample separately for the two types of BGs. Specifically, we let μ_s and η_s with $s = M, D$ be the parameters capturing the different advising needs of subsidiaries in multinational and domestic BGs, respectively. Then, the advising time constraint becomes $n_{\ell-1}e^{\tau} = n_{\ell}e^{(\varphi+\eta_s-\theta)\ell+\mu_s}$. By interaction, this expression gives $n_{\ell} = e^{\ell\tau_s+\frac{1}{2}\ell(\ell+1)(\theta-\varphi_s)}$ with $\tau_s = \tau - \mu_s$ and $\varphi_s = \varphi + \eta_s$, which we run as a log-regression as before. The values obtained are $\tau_M = 0.851$ and $\varphi_M = 0.169$ for multinational BGs, and $\tau_D = 1.453$ and $\varphi_D = 0.646$ for domestic ones. Hence, we have $\tau_M < \tau < \tau_D$ and $\varphi_M < \varphi < \varphi_D$. These estimates suggest that, while multinational BGs exhibit more effective communication, their subsidiaries demand more advising time to solve their problems.

What the alternative estimates imply is depicted in Figure 9, which again considers a representative US BG in the top 95 percentile of the cutoff layer distribution. Adjusting the BG's production efficiency to hold the cutoff layer fixed at $\ell^* = 6$, the figure simulates how its hierarchical structure differs when it is assigned the advising and communication efficiencies estimated in the pooled sample (τ and φ; gray bars), in the multinational sub-sample (τ_M and φ_M; black bars), and in the domestic sub-sample (τ_D and φ_D; white bars). The figure shows that, for given production efficiency, whereas the multinational hierarchy is left skewed as the pooled one, the domestic hierarchy is right skewed. This happens for two reasons. On the one hand, the number of subsidiaries per layer n_ℓ reaches the top of its hump-shaped relation for smaller ℓ in the domestic case as $\tau_D/(\varphi_D - \theta) < \tau_M/(\varphi_M - \theta)$ holds. On the other hand, in order to have six layers, profits have to be higher when n_ℓ reaches the top in the domestic case as communication gets more

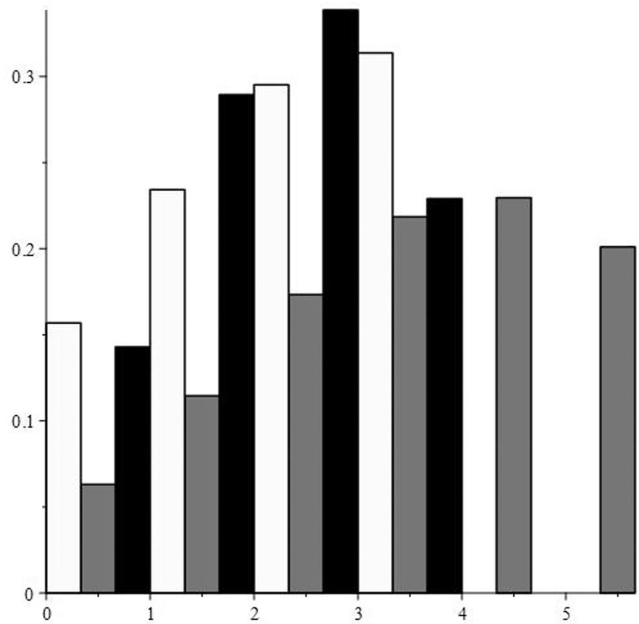

Figure 10. Hierarchies are less left-skewed in the USA and Japan than in Europe. Model simulation. Density of subsidiaries across layers for BGs with $\omega = 2.131$ in C. Europe (white), USA (gray) and Japan (black). Parameter values: $A = F = w = 1$, $\theta = 0.0207$, $\sigma = 5$, $\tau = 0.742$ and $\varphi = 0.191$ (white), $\tau = 0.961$ and $\varphi = 0.203$ (gray), $\tau = 1.802$, and $\varphi = 0.569$ (gray).

rapidly harder as n_ℓ keeps increasing. Yet, within domestic and within multinational BGs, it is still true that those organized in a larger number of hierarchical layers tend to have more subsidiaries and are less likely to be organized as pyramids.

Next, to show that the model correctly predicts that left-skewed hierarchies are less frequent in the USA and Japan than in Europe, we consider a US BG in the top 95 percentile of the cutoff layer distribution, and ask how its hierarchical structure differs from a European or a Japanese BGs with the same production efficiency. This requires using the model to simulate the hierarchy arising from keeping production efficiency constant at $\omega = 2.131$ while switching advising efficiency and communication efficiency from the US estimates ($\tau = 0.961$, $\varphi = 0.203$) to the European ($\tau = 0.742$, $\varphi = 0.191$) or Japanese ones ($\tau = 1.802$, $\varphi = 0.569$) as reported in Table 7. Figure 10 depicts the density distribution of subsidiaries across layers for the three countries. The gray bars correspond to the US BG and are the same as in Figure 8, while the black and white bars belong to the Japanese and European BGs, respectively. The figure then shows that the model predicts that the US and Japanese distributions are less left-skewed than the European one, which is what we see in the data.

Based on the discussion in Section 4.3, such heterogeneity in hierarchical structures can be explained by the combination of two factors. First, n_ℓ reaches the top of its hump-shaped relation with ℓ at a lower layer (i.e., larger ℓ) in the US than the European BGs, and at a lower layer in the European than in the Japanese BGs. This is due to the fact that $\tau/(\varphi - \theta)$ equals 5.271, 4.357 and 3.286 for the US, European, and Japanese BGs, respectively. Second, for the same productivity $\omega = 2.131$, the US and Japanese BGs are able to go beyond their humps' tops, while EU is not. This is due to the fact that the US and Japanese

BGs generate more profits at each layer thanks to the larger numbers of subsidiaries their superior advising efficiency allows for.

5. CONCLUSIONS

Hierarchical differentiation is a cornerstone of the organizing process. While research has traditionally focused primarily on pyramid-shaped hierarchies, alternative hierarchical shapes may be at least as relevant. In this article, we have done three things. First, exploiting a newly assembled dataset, we have provided the first worldwide overview of the patterns of hierarchical differentiation across BGs, highlighting the coexistence of different hierarchical shapes. Second, we have shown how the different shapes can arise as optimal hierarchical structures in a parsimonious knowledge-based model of BGs when subsidiaries operations involve ubiquitous problem solving under parents' supervision. Third, we have checked the consistency of some of its key predictions with the empirical patterns observed in our dataset.

APPENDIX A: OWNERSHIP DATA AND ROBUSTNESS

In this Appendix, we provide some more information on the basic ownership data and how we use them for the purpose of our analyses. Moreover, we describe more in detail the characteristics and the differences in the determination of BGs that can be retrieved from Bureau Van Djik data either following the approach proposed by Sonno (2025) or by Rungi et al. (2017).

A.1 Construction of BGs

To construct the BGs for our analysis, we first identify the ISH of each subsidiary and then assign hierarchical layer 0 to the parent (GUO), hierarchical layer 1 to the subsidiaries directly controlled by the parent (i.e., one degree of separation), and analogously layer $\ell+1$ to the subsidiaries with ISH at generic layer ℓ (i.e., ℓ degrees of separation). When a subsidiary is controlled through one or more subsidiaries, we identify it as indirectly controlled by the parent. Layers can be interpreted as a measure of hierarchical distance from the HQ: the higher the number assigned to each layer, the higher the hierarchical distance. The number assigned to the most distant layer L defines the depth of the hierarchy with layers indexed $\ell = 0, \ldots, L$. Figure A1 provides a simple illustrative example of the construction of the hierarchy of a BG starting from the ownership links. It identifies A

Subsidiary	Shareholder	Relation
B	A	GUO
B	A	ISH
C	A	GUO
C	B	ISH

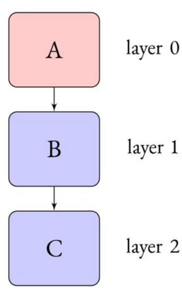

Figure A1. From ownership links to BGs. The table on the left shows a simplified example of four links in the Orbis Historical Ownership Database. From the links we can reconstruct the group in the diagram on the right. We observe that A is the parent (GUO) of both B and C, and that A is the direct owner (ISH) of B while B is the direct owner of C. Therefore, we can conclude that A owns C indirectly through B.

as a parent company because the company is reported both as a GUO and as the ISH of company B, which in turn can be located at hierarchical layer $\ell = 1$. Subsidiary C, instead, can be located at layer $\ell = 2$ of the BG with company A as parent given that its ISH is a company located at layer 1 in that BG. The depth of the BG's hierarchy is $L = 2$.

A.2 An alternative construction of BGs

These firm-level databases collect original information from a variety of national and international registries, regulatory bodies, companies' annual reports, websites and specialized press. For our purpose, we extract information on shareholding activity for companies active in more than 200 countries in the year 2015. The starting point of the two approaches is different, however, how we will see in this section, they both retrieve a comparable dataset of BGs worldwide.

In the article, we follow Sonno (2025), who starts from the Historical Ownership Database by Bureau Van Dijk. This dataset provides for each company information on all its shareholders and identifies several types of relations. The author proposes an algorithm that retrieves the hierarchical distance of a company from its parent company using two types of relations: the corporate GUO with at least 50.01% of voting rights (GUO 50C hereafter), that is the highest corporate independent shareholder in the shareholding structure of a company, and the ISH, that is the first shareholder in the path from an affiliate to its GUO. Combining the definition of GUO 50C and ISH, with the fact that each shareholder is reported more than once depending on its role in the shareholding structure of a subsidiary, it is possible to create a routine that counts the steps leading a subsidiary to its parent. Throughout this approach we rely on the definition of direct or indirect majority (50.01%) of voting rights provided by Bureau Van Djik. Therefore, we use an exogenous definition of control. Using this procedure, it is possible to obtain a dataset of 2,901,466 parent companies controlling 5,676,289 subsidiaries for 2015, as we have excluded branches from our analysis. Also note that as we require additional details on the type of ownership links, the latter yields different subsamples of the dataset depending on the information available to us. As an example, to document the sector of parents and affiliates (allocated in a specific layer within the BG) we rely on a sample of 2.2M parents and 4.6M affiliates. A detailed breakdown of the different sample compositions can be found in Table 1.

We test the robustness of our results after replicating the methodology proposed by Rungi et al. (2017). The authors use data from the Ownership section of the Orbis Database, where Orbis collects all the original information available on shareholdings. For each company, we have a list of all (individual, corporate or state) shareholders.

Any time a company invests in the equity of another company, an ownership network is generated such that voting rights can be separated from cash rights. In modern economies, corporate ownership structures can become fragmented (see, e.g., La Porta et al. 1999), and the identification of ultimate parent companies can become difficult, especially in the case of MNEs crossing national borders UNCTAD (2016). The authors model a backward solution for a voting rule across interlocking assemblies of shareholders. Assemblies of shareholders interlock when (individual and corporate) shareholders generate cross-holdings, ownership cycles and multiple ownership paths. When shareholding activity interlocks, a coordination effort is required across different ownership paths to enforce management decisions starting from HQs. Bottom-up, this approach extracts hierarchies of firms made of parent companies and their subsidiaries ordered on hierarchical layers by considering all the ownership paths that can connect any two nodes (companies/shareholders) in an ownership space. Starting from a basic ownership matrix including all the shareholding links between companies and shareholders, Rungi et al. (2017) iteratively simulate corporate control after assuming that the latter entails cases of: (i) *direct control*, when the parent company holds the absolute majority of voting rights in a subsidiary; (ii) *indirect control by transitivity*, when the parent company has direct control of a subsidiary that in turn has direct control over another subsidiary, in a

Table A1. Geographic coverage—robustness.

Region or country	BGs All	%	BGs Multinational	%	Subsidiaries All	%	Subsidiaries Foreign	%
Africa	5102	0.22	4169	2.07	30,346	0.64	17,088	2.27
Asia	105,449	4.45	19,142	9.51	316,014	6.67	99,624	13.24
Australia	58,788	2.48	2771	1.38	136,189	2.87	14,750	1.96
EU 28	600,829	25.35	111,522	55.41	1,625,508	34.29	387,006	51.44
Other Europe	36,073	1.52	14,089	7.00	84,045	1.77	22,441	2.98
Latin America	30,058	1.27	18,247	9.07	83,227	1.76	51,693	6.87
Russia	29,741	1.25	974	0.48	110,232	2.33	50,541	6.72
USA	1,435,218	60.56	22,511	11.18	2,138,025	45.10	63,220	8.40
Rest of the world	68,634	2.90	7847	3.90	216,766	4.57	45,992	6.11
Total	2,369,892	100.00	201,272	100.00	4,740,352	100.00	752,355	100.00

Notes: The table refers to the dataset constructed following Rungi et al. (2017). It details the number of observations of subsidiaries and BGs in a list of regions or countries. It also specifies the number of BGs that control at least one subsidiary abroad (multinational), and that of subsidiaries that are controlled by foreign parents. For regional aggregation, we have used the online version of the United Nations publication Standard Country or Area Codes for Statistical Use.

Table A2. Hierarchical distance—robustness.

Layer	Domestic Subsidiaries	%	Foreign Subsidiaries	%	All Subsidiaries	%
1	1,579,504	73.51	228,542	36.11	1,808,046	65.00
2	376,511	17.52	186,062	29.40	562,573	20.22
3	123,021	5.73	104,063	16.44	227,084	8.16
4	43,079	2.00	55,414	8.76	98,493	3.54
5	15,354	0.71	28,135	4.45	43,489	1.56
6	5934	0.28	14,182	2.24	20,116	0.72
7	2518	0.12	8132	1.28	10,650	0.38
8	1321	0.06	3765	0.59	5086	0.18
9	600	0.03	2104	0.33	2704	0.10
10	268	0.01	987	0.16	1255	0.05
>10	580	0.03	1552	0.25	2132	0.08
Total	2,148,690	100.00	632,938	100.00	2,781,628	100.00

Notes: The table refers to the dataset constructed following Rungi et al. (2017). It shows the number of subsidiaries per layer, distinguishing between domestic and foreign subsidiaries. It excludes BGs with only one subsidiary.

sequence; (iii) *indirect control by consolidation* of voting rights, when a parent company can control a subsidiary by summing up to a majority of capital shares held in her portfolio and in the portfolio of other subsidiaries; (iv) *dominant shareholding* when a parent company can control a company with minority stakes because other minority shareholders are too fragmented to form an opposing coalition. For the purpose of this article, we limit our analyses only to the first three cases, excluding control by dominant minorities, although all of them find a correspondence in international accounting standards (OECD 2005, 2008; UNCTAD 2009; IFRS 2011).

As we can see by comparing relevant descriptive statistics, the results obtained with the two approaches are highly consistent with each other. Table A1, for example, shows the geographic coverage of the data obtained following Rungi et al. (2017). Although the figures are slightly lower than those obtained following Sonno (2025) and displayed in Table 2, the correlation between the two tables is over 0.99 for every category considered, even if we break down the table further into

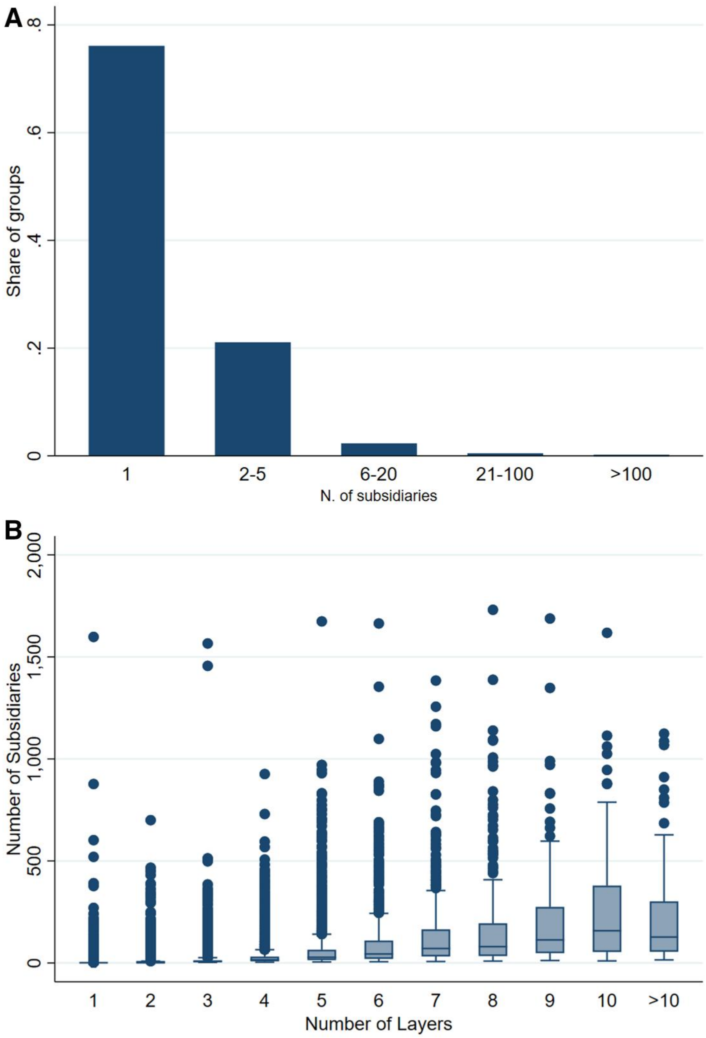

Figure A2. Hierarchical description—robustness. The graph refers to the dataset constructed following Rungi et al. (2017). Panel A shows the distribution of groups' dimension in terms of number of subsidiaries. Panel B presents eleven box-plots of groups' dimension conditional on the number of layers their hierarchies display. We have excluded the 14 BGs with more than 2000 subsidiaries so as to make the figure more legible.

21 countries and regions.[23] Similarly, looking at Tables A2 and 3 that detail the distribution of domestic and foreign subsidiaries among hierarchical layers, we observe slightly different numbers but very similar shares. Again, the figures in the two tables correlate at more than 0.99. Finally, looking at the Panel A of Figure A2 we observe that the distribution of groups in terms of the subsidiaries

[23] In following the approach proposed by Sonno (2025), we use the 2019 release of the Historical Ownership Database of Bureau van Dijk, while the dataset elaborated by Rungi et al. (2017) uses the 2016 release of the Ownership section of the Orbis database by Bureau van Dijk. This discrepancy partially explains the higher number of observations found by the Sonno (2025) approach.

Table A3. Number of affiliates per layer across BGs—robustness.

	Maximum layer of the BG:										
Layer	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**>10**
1	1.3	3.2	7.3	16.3	39.3	60.5	86.1	75.6	80.3	91.8	118.6
2		2.6	7.1	15.6	27.2	37.2	49.0	97.9	71.3	89.3	78.6
3			3.5	9.9	21.3	42.1	45.0	77.0	168.7	72.1	145.2
4				4.2	11.3	23.1	35.1	32.4	82.0	92.4	79.7
5					4.0	11.4	21.2	23.4	35.5	36.0	32.4
6						5.2	13.3	25.7	34.3	55.0	28.9
7							5.5	13.1	21.7	39.7	58.8
8								3.5	12.6	22.8	19.6
9									4.7	9.8	18.4
10										2.5	15.2
>10											10.7
N. of BGs	2,367,786	122,697	24,609	6829	2491	1108	561	278	150	83	117

Notes: The table refers to the dataset constructed following Rungi et al. (2017). It shows the average number of subsidiaries per layer for BGs having different maximum layers. If BGs with only one subsidiary were excluded, the mean number of affiliates in the first layer for BGs with only one layer would be 2.8.

they control is extremely skewed, with more than 75% of the groups controlling only one subsidiary, and only 0.08% of them having more than 100 subsidiaries. Once more, this is very similar to what we observed in Figure 1 using Sonno (2025)'s approach.

One feature fundamental to our analysis is the hierarchical structure of BGs. To assess any potential difference between the two samples in this respect, we look at Tables 4 and A3, that refer, respectively, to the samples following Sonno (2025) and Rungi et al. (2017), and report on the average composition of BGs, conditional on the number of layers of their hierarchy. Even if the two tables are qualitatively very similar, and both distinctly show groups shaped as "inverted pyramids" (groups with more subsidiaries in the first layers and fewer in the lowest ones), it is worth noticing one main difference: the dataset obtained following Sonno (2025) show a slightly higher number of subsidiaries in the first hierarchical layer. This is because the approach of Rungi et al. (2017) specifically catches *cross shareholding*, thus potentially detecting, in the case of fragmented shareholding, additional layers to BG structures of any size. Indeed, we detect a decrease of BGs with at most one layer in favor of the other categories.

To test the robustness of our findings on the relationship between routinizability and the hierarchical position of subsidiaries, Figure A3 replicates Figure 4 using the dataset constructed based on the approach proposed by Rungi et al. (2017). This replication confirms the consistency of our original results with this alternative dataset.

APPENDIX B: STYLIZED FACTS AND ROBUSTNESS

B.1 Real-world cases

The following section introduces three real-world examples of BGs characterized by distinct organizational hierarchies by means of visualization with the GEPHI program (see also Rungi et al. (2017)). Figure B1 provides a stylized representation of the hierarchical structure of a pyramid-shaped BG, together with the skewness of the distribution of its subsidiaries across layers. Similarly, Figures B2 and B3 present analogous information for a diamond-shaped and an inverse-pyramid-shaped BG, respectively. The skewness values for the three examples align with the literature (Wellman et al. 2020). Observing Figure B1, in the case of a pyramid-shaped hierarchy, the skewness appears to be negative (−0.105). Conversely, for an

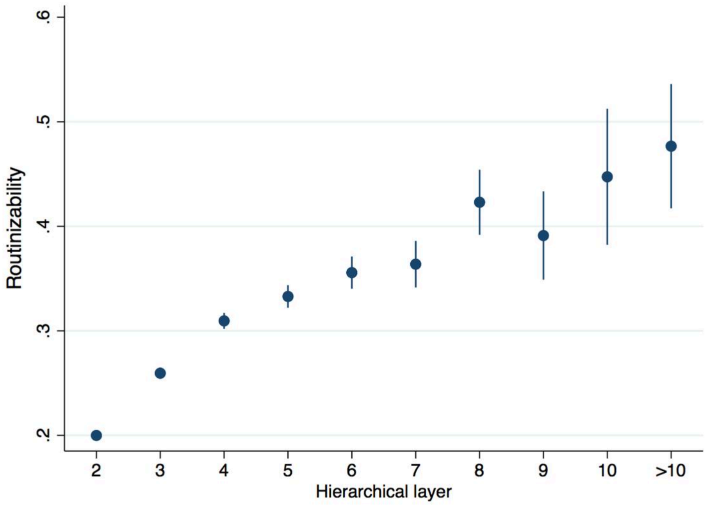

Figure A3. Routinizability over hierarchies—Rungi et al. (2017). The graph shows the coefficients and 95% confidence intervals obtained from a regression of the routinizability index of each subsidiary on the hierarchical layer of the same subsidiaries, including group and host country FE, and using robust standard errors using the dataset elaborated by Rungi et al. (2017). Layer 1 constitutes the omitted category.

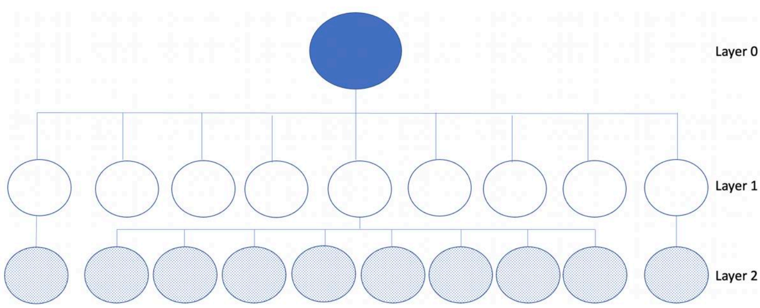

Figure B1. Pyramid-shaped BG. The figure shows a visual representation of the hierarchical structure of a pyramid-shaped BG, together with a table including relevant descriptive statistics. Average outdegree centrality corresponds to the mean of outdegree centrality within the hierarchy, computed disregarding layer 0 of the HQ.

inverse-pyramid-shaped hierarchy, the skewness seems to be positive (1.336), as depicted in Figure B3. Finally, Figure B2 shows that the skewness for a diamond-shaped BG is closer to zero (0.573) compared to the two alternative cases. Furthermore, the three figures provide information on the average outdegree centrality of subsidiaries within these different hierarchical structures. This metric reflects the average number of subsidiaries controlled by any other subsidiary within the hierarchy. This measure is relevant as it serves as a control in the robustness checks for the stylized facts presented in the subsequent figures.

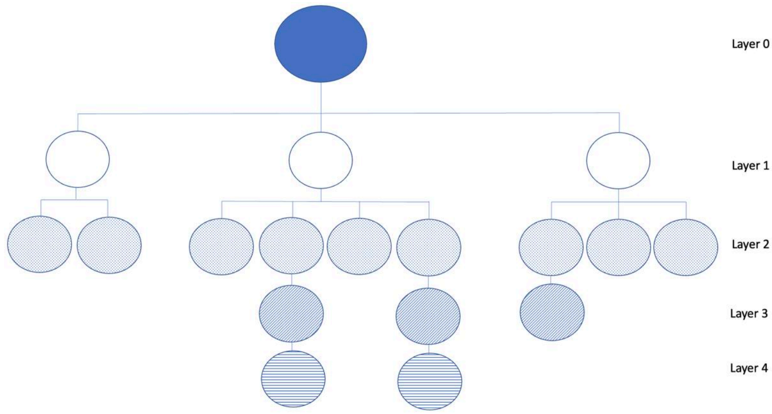

Figure B2. Diamond-shaped BG. The figure shows a visual representation of the hierarchical structure of a diamond-shaped BG, together with a table including relevant descriptive statistics. Average outdegree centrality corresponds to the mean of outdegree centrality within the hierarchy, computed disregarding layer 0 of the HQ.

B.2 Robustness

Here we present additional tables and figures that test the robustness of our results along several dimensions, together with the regression behind our graphs in Section 3.

B.2.1 Including ownership links in the analysis

It could be argued that abstracting from actual ownership relationships within BGs might lead to omitting a relevant aspect of hierarchical structures in our empirical analysis. In Section 3.4, we referenced management literature that supports our approach, drawing on a substantial body of research that provides evidence of knowledge transfers within BG structures, independent of ownership links (e.g., Gupta and Govindarajan 1991; Bresman et al. 1999; Ciabuschi et al. 2011; Crespo et al. 2020).

In this section, we test the robustness of our empirical analysis by directly accounting for the actual ownership links among subsidiaries. We first present some statistics on the average number of owned subsidiaries per owner subsidiary (among the subsidiaries that own some other subsidiaries), computed for each BG. 54% of BGs in our sample consist of groups where each subsidiary at layer $\ell - 1$ is linked to only one subsidiary at lower layer ℓ. For these groups, the average is, by construction, equal to 1. In contrast, for the remaining 46% of BGs, the statistics is greater than 1, with an average of 2.99 owned subsidiaries per owner subsidiary and a standard deviation of 1.68. In particular, subsidiaries belonging to the 46% of groups with an average number of owned subsidiaries per owner subsidiary larger than 1 account for 80% of all subsidiaries.

Based on these preliminary data, we tested the robustness of our results by performing a number of additional exercises.

For the purposes of this discussion, it is important to emphasize that the HQs is excluded from the computation of skewness when analyzing the structure of BGs. This exclusion is due to the legal requirement that a BG must have only one "ultimate owner." Consequently, when calculating skewness, the count of subsidiaries by layer always begins at layer 1, which is the first layer where the organizational structure can be chosen. We have built a robustness exercise by specifically focusing on the issue of "dead branches" in hierarchical structures. Indeed, when including actual ownership links in the analysis, the presence of dead branches can affect the

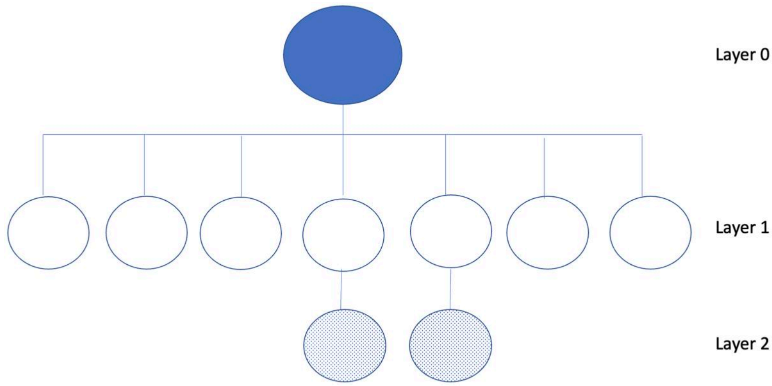

Figure B3. Inverse-pyramid-shaped BG. The figure shows a visual representation of the hierarchical structure of an inverse-pyramid-shaped BG, together with a table including relevant descriptive statistics. Average outdegree centrality corresponds to the mean of outdegree centrality within the hierarchy, computed disregarding layer 0 of the HQ.

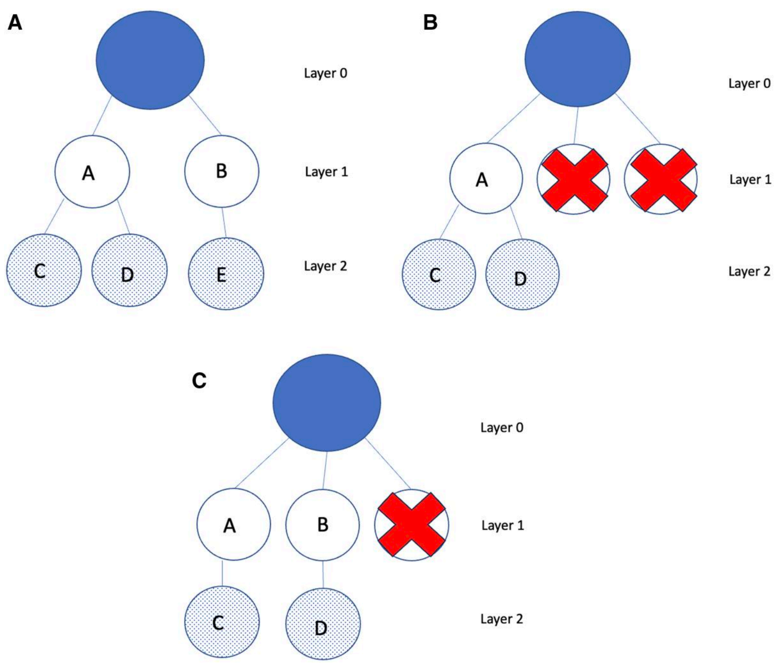

Figure B4. Hierarchical structures with and without dead branches.

skewness of a BG's hierarchy. Let us clarify this statement considering the example from Figure B4.

In this example, Case 2 and Case 3 would result in identical skewness when counting the number of subsidiaries across layers, as it is done in our analysis. However, when considering only subsidiaries that own other subsidiaries, there would be two measures of skewness.

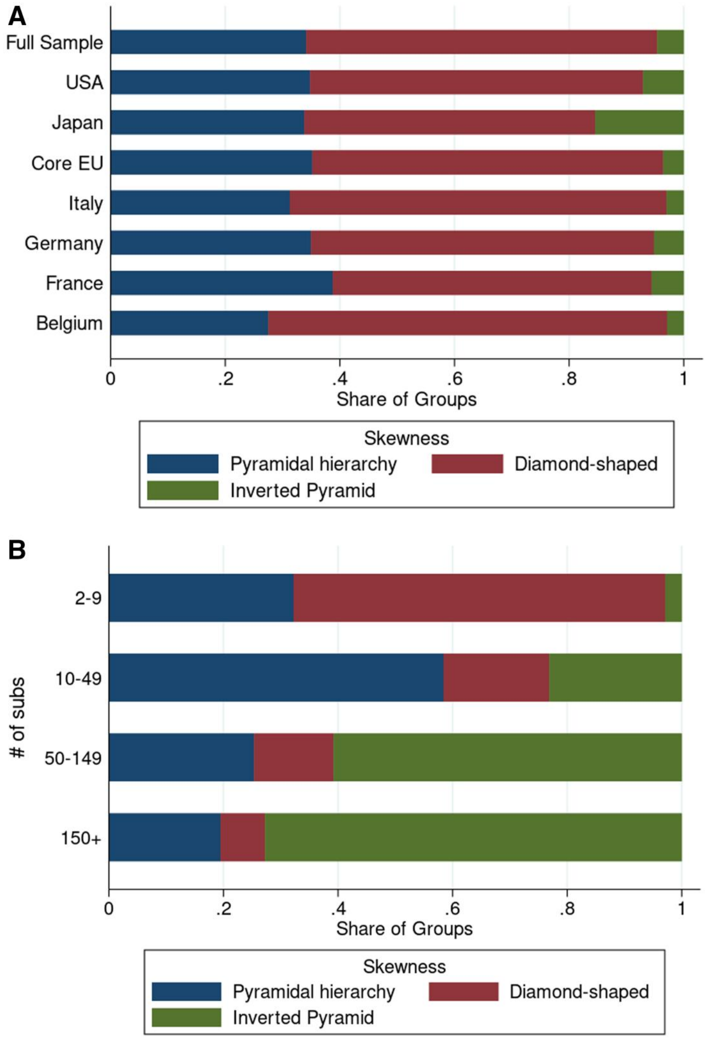

Figure B5. Skewness bars—skewness clean metric. The figure shows the frequency of the three types of hierarchical structures. In Panel A, BGs are grouped by geographical area, while in Panel B, they are grouped by size. BGs with layers less than 2 are excluded because their skewness can only take value 0 and thus alter the graphs. We have removed subsidiaries that do not control any other subsidiary (i.e., outdegree centrality = 0) and are not in the last layer of the BG's hierarchy.

To address this, the main stylized facts presented in Section 3 have been revisited, employing two different approaches based on the outdegree centrality of subsidiaries within the group hierarchy. First, we introduced a "skewness clean" metric, where the skewness of BGs' hierarchies is calculated excluding dead branches. Then, we repeated stylized facts adding the outdegree centrality of subsidiaries as a control variable. Figure B5 replicates Figure 3 excluding dead branches, showing that the distribution of BGs across skewness levels and the prevalence of the three types of hierarchical structures for this restricted sample are comparable to those observed in the complete one.

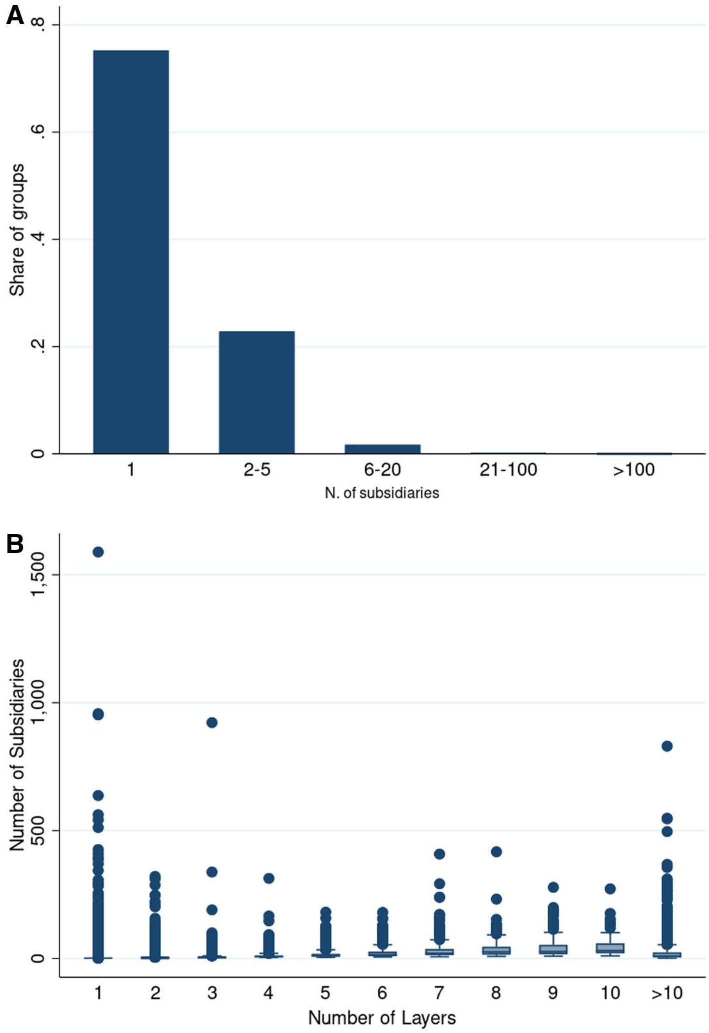

Figure B6. Hierarchical description—skewness clean metric. Panel A shows the distribution of groups' dimension in terms of number of subsidiaries. Panel B presents eleven box-plots of groups' dimension conditional on the number of layers their hierarchies display. We have removed subsidiaries that do not control any other subsidiary (i.e., outdegree centrality = 0) and are not in the last layer of the BG's hierarchy.

Similarly, Figure B6 and Table B1 replicate Figure 1 and Table 4 excluding dead branches from the sample, once again presenting consistent results.

Finally, Figure B7 examines the relationship between routinizability and the hierarchical position of subsidiaries, addressing the issue of dead branches using two different approaches. Panel A replicates the analysis from Figure 4, excluding dead branches. Panel B replicates the analysis from Figure 4 for the full sample but adding the subsidiary outdegree centrality as a control in the regression. Results are once again entirely consistent with our analysis in Figure 4.

Table B1. Number of affiliates per layer across BGs—skewness clean metric.

	Maximum layer of the BG										
Avg. Subs per layer	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**> 10**
1	1.37	1.15	1.70	2.50	3.84	5.45	6.89	7.53	8.29	9.60	13.29
2		2.18	1.35	2.44	3.93	5.37	6.82	8.60	9.73	8.13	12.95
3			2.80	1.49	3.01	4.53	6.51	7.53	8.49	7.77	12.35
4				3.06	1.60	3.06	4.77	5.93	7.64	5.75	9.67
5					3.32	1.63	3.19	4.26	5.29	4.71	7.86
6						3.36	1.62	3.01	4.15	4.19	6.81
7							3.80	1.53	2.92	3.60	5.94
8								3.05	1.57	2.46	4.55
9									3.40	1.43	4.47
10										2.64	2.93
> 10											8.52
N. of BG	2,739,149	126,168	24,172	6811	2571	1157	618	350	207	91	138

Notes: The table shows the average number of subsidiaries per layer for BGs having different maximum layers. We have removed subsidiaries that do not control any other subsidiary (i.e., outdegree centrality = 0) and are not in the last layer of the BG's hierarchy. If BGs with only one subsidiary were excluded, the mean number of affiliates in the first layer for BGs with only one layer would be 2.81.

B.2.2 Controlling for size of BG and subsidiaries

To test the robustness of our results, we replicated the findings on routinizability from Figure 4 using size controls for both BGs and subsidiaries, with both consolidated and unconsolidated balance sheet data.

Specifically, Panel A of Figure B8 replicates Figure 4 by splitting the BGs into quartiles based on consolidated balance sheet data. To classify BGs into quartiles, we used the distribution of total assets. When total assets were unavailable, we used (in order) operating revenues, the number of employees, and capital as alternatives. This approach was implemented to address the data limitations encountered when consolidated balance sheet information was included in the analysis. For the same reason, we restricted this analysis to BGs with at least four layers and displayed only up to eight layers instead of ten, as the final layers were notably underpopulated. Panel B of Figure B8, on the other hand, replicates Figure 4 by incorporating a set of dummies indicating the quantiles of subsidiary size, which were determined using the same approach as for BGs, but based on unconsolidated balance sheet data. Results are consistent with those from our main analysis.

B.2.3 Separating MNEs and domestic BGs

As a final robustness check, we replicated the results on routinizability shown in Figure 4 separately for MNEs and domestic BGs, and we added a control for the geographic distance of subsidiaries from their HQs. The results are presented in Figure B9. The graphs show the coefficient estimates from the regression of routinizability on the hierarchical layer and BG fixed effects. Panel A presents the results for domestic subsidiaries. Since the sample is restricted to domestic BGs, we do not include subsidiaries' country fixed effects, as they would be absorbed by the BG fixed effects. In Panel B, the sample is restricted to multinational BGs, and we control for subsidiaries' country fixed effects. Finally, in Panel C, we also focus on multinational BGs but include an additional control for the distance between the subsidiaries' countries and the HQs' ones, calculated as the distance between capital cities. Reassuringly, our findings remain robust across all these variations.

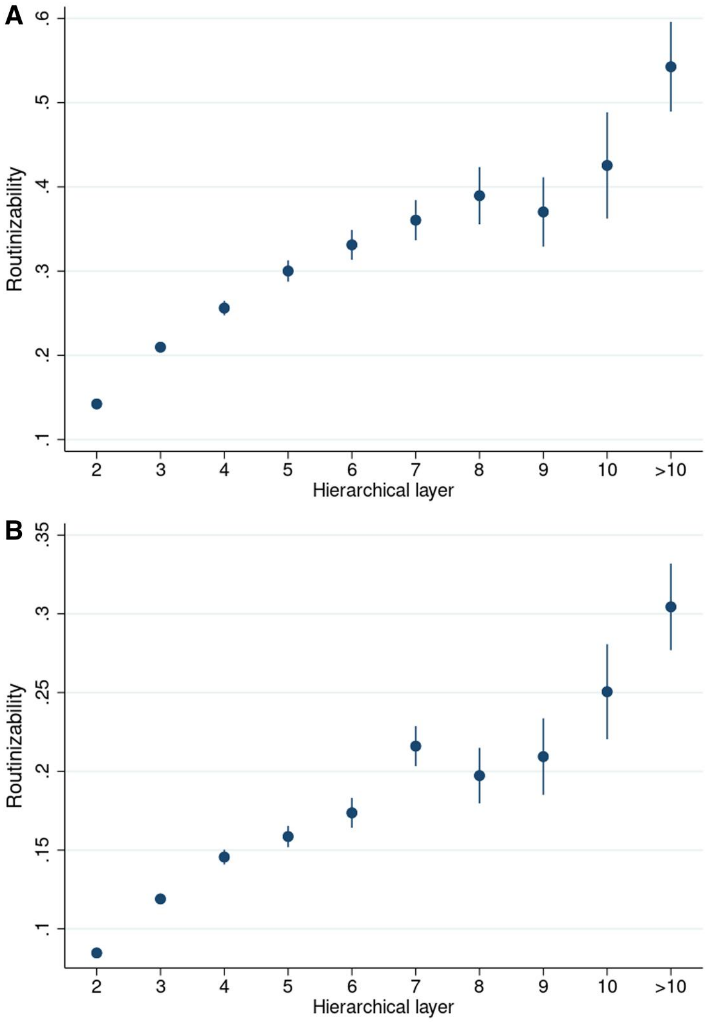

Figure B7. Routinizability over hierarchies—misclassification concerns. Subtitle Panel A: skewness clean metric. Subtitle Panel B: controlling for outdegree centrality. The two graphs show the coefficients and 95% confidence intervals obtained from a regression of the routinizability index of each subsidiary on the hierarchical layer of the same subsidiary, including group and host country FE, and using robust standard errors. In Panel A, the subsidiaries that do not control any other subsidiary (i.e., outdegree centrality = 0) and are not in the last layer of the BG's hierarchy are removed from the estimation sample. In Panel B, the outdegree centrality of the subsidiary is included as a control.

B.3 Additional exhibits

In this section, Table B2 presents the regression underlying the analysis in Figure 4.

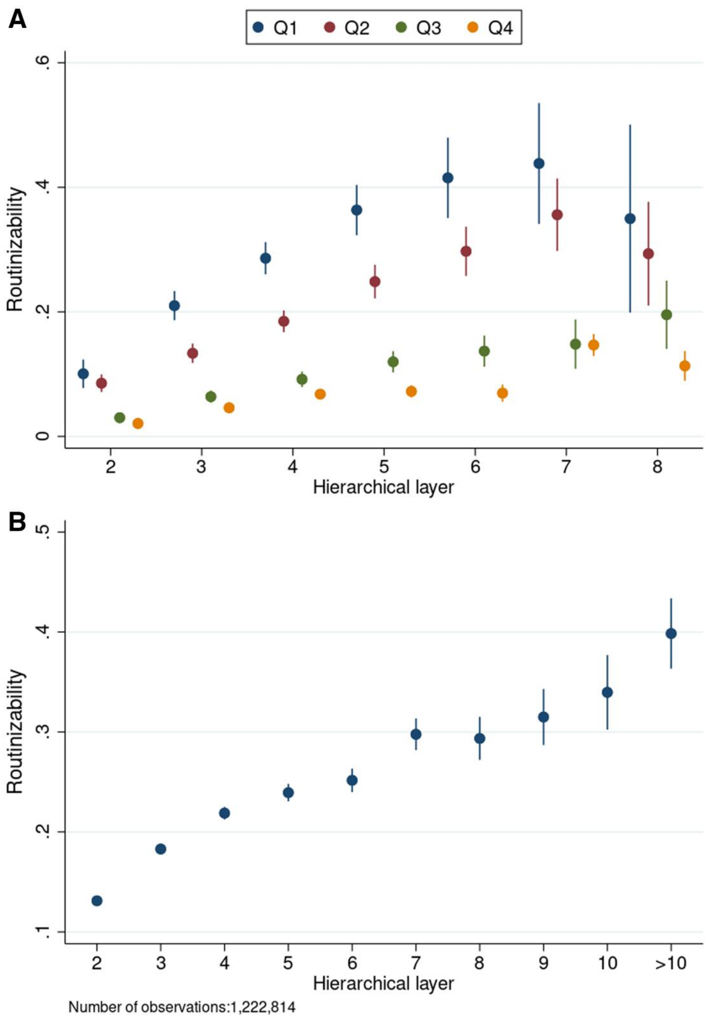

Figure B8. Routinizability over hierarchies—misclassification concerns. Subtitle Panel A: BG—consolidated. Subtitle Panel B: subsidiary—unconsolidated. Panel A shows coefficients from regression of routinizability on hierarchical layer for different BG size samples. The dataset is restricted to BGs with at least 4 layers and the sample is divided into quartiles of consolidated BG size measured as quartiles of *total assets, operating revenue, number of employees, capital* (if one variable is missing, we use the next variable in this order to limit the number of missing values). Some high-level layers are underpopulated especially for the first two quartiles. Therefore, we show the first 8 layers. Panel B shows the coefficients of the regression of routinizability on hierarchical layer dummies and a set of dummies for quartiles of subsidiary size.

APPENDIX C: TRANSMISSIONS ALONG THE HIERARCHY

In this section, we corroborate the findings on knowledge transmission within BGs (section 3.3) by showing, first, that self-citations rise as we move down the hierarchy, and second, by

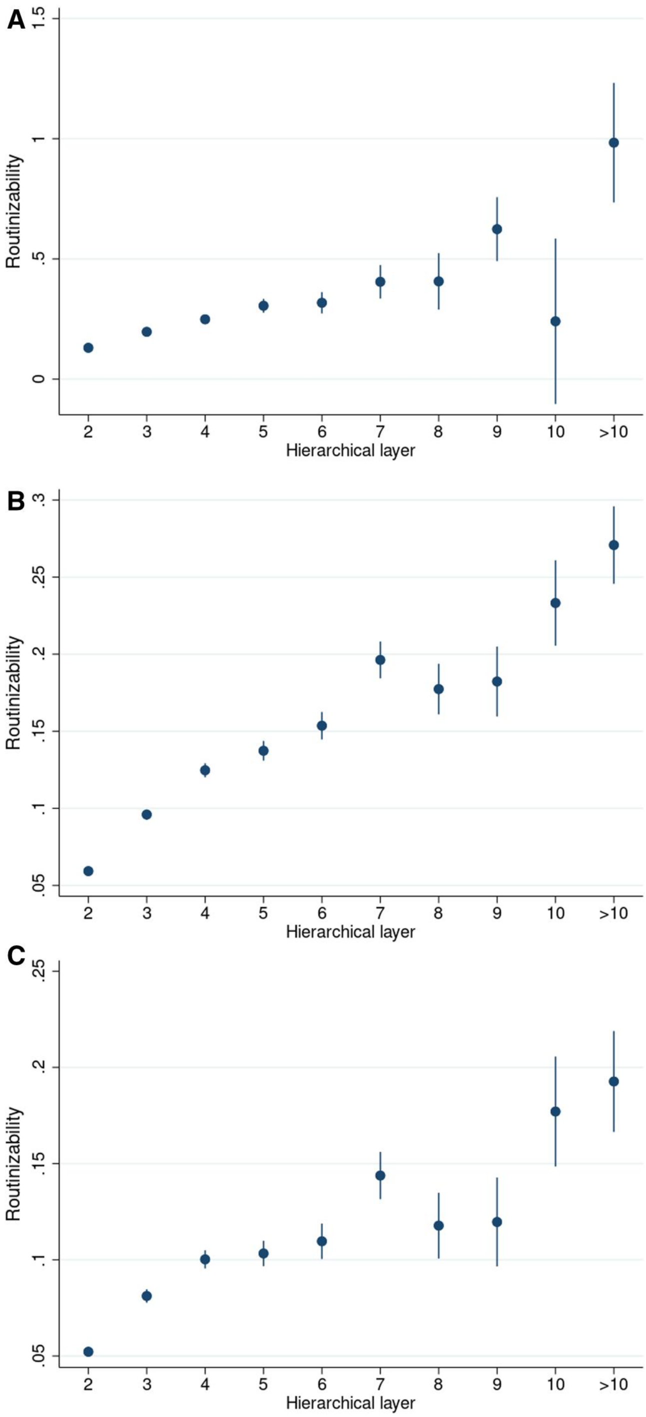

Figure B9. Routinizability over hierarchies—MNE vs domestic. Subtitle Panel A: domestic. Subtitle Panel B: MNE, FE. Subtitle Panel C: MNE, distance. These graphs show the coefficient estimates of the regression of routinizability over the hierarchical layer and BG FE. In Panel A, we

Table B2. Routinizability over hierarchies—coefficients.

Dependent variable	**Routinizability**
Estimation method	OLS
Layer 2	0.088***
	(0.001)
Layer 3	0.123***
	(0.002)
Layer 4	0.151***
	(0.002)
Layer 5	0.164***
	(0.003)
Layer 6	0.179***
	(0.005)
Layer 7	0.222***
	(0.006)
Layer 8	0.203***
	(0.009)
Layer 9	0.215***
	(0.012)
Layer 10	0.256***
	(0.015)
Layer >10	0.311***
	(0.014)
Parent FE	Yes
Country FE	Yes
Observations	2,687,600
R^2	0.568

Notes: OLS estimations. *Routinizability* is the routinizability of the layer's activity, as retrieved from the PDII at the three-digit NAICS 2002 sector level; *Layer* is the degree of separation between the HQ and the corresponding subsidiary. The constant is omitted from the table. All specifications include parent company FE and subsidiary's country FE. The model is estimated on the full world sample, excluding BGs with only one affiliate. Heteroskedasticity robust standard errors at the HQ level are in parentheses.

*** $p < .01$.

presenting evidence of a positive and significant correlation between productivity shocks in the lower layers and productivity changes at the HQs.

C.1 Citations within BGs

In this section we exploit the "quality" dimension of patenting activity by looking at citations. We collapse our firm-patent citation data to calculate the average number of citations per firm at each hierarchical level, differentiating between self-citations and external citations. Figure C1 displays the average proportion of self-citations to total citations across layers. This proportion grows as we move down the hierarchy, suggesting that higher-quality patents (those receiving external citations) are more likely to be found in layers closer to the HQ, while lower layers primarily rely on self-citations.

Figure B9. Continued
cannot add the subsidiaries' country FE, because we restrict the sample to domestic BG only. In Panel B, we restrict the sample to MNE BG and control for subsidiaries' country FE. In Panel C, we restrict the sample to MNE BG and add a control for the distance of subsidiary countries from the HQ country. Distance is computed as the distance between the capital cities.

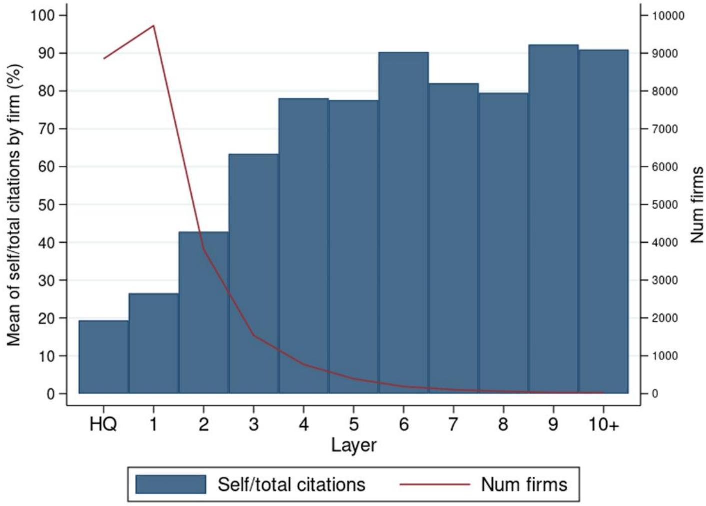

Figure C1. Specialization of inventing activities: self-citations/total citations. The bar chart shows the ratio between the average number of patent self-citations made by a firm and the average number of patent citations that a firm receives in general (from all other firms in the dataset). This ratio is shown as a percentage on the left y-axis. It should be noticed that higher hierarchical layers have very few subsidiaries, both because there are fewer observations per layer and because there are less BGs that have so many layers. The continuous line shows the number of subsidiaries we are considering at each layer (see right y-axis).

C.2 Production efficiency transmission within BGs

In this section, we discuss whether production efficiency shocks are transmitted from higher to lower layers to provide evidence of interactions along the hierarchy as highlighted by our model. For this exercise, we produce a panel version of our sample (see Section 2). Specifically, we replicate the procedure followed to obtain the 2015 worldwide cross-section of BGs from 2007 to 2014.[24] We then match BG data over time to balance sheet information for HQs and subsidiaries retrieving labor productivity as a proxy for production efficiency.[25] On average across years our baseline estimation sample consists of 2,760 groups (2729 of which with more than one subsidiary) and 36,235 subsidiaries in total. Given better coverage of financial data in Europe, EU groups represent more than 86% of the entire sample (47% of which are MNEs).

In both panels of Table C1, we regress the mean yearly change in labor productivity of the subsidiaries in the three bottom layers on the mean yearly change in labor productivity of the subsidiaries in the first layer. Panel A includes all BGs, while Panel B focuses only on BGs with at least four layers. In column 1 we do not include any fixed effect, in column 2 we include HQ fixed effects, in column 3 we include HQ's country fixed effects, in column 4 year fixed effects, and in column 5 both country and year fixed effects. All OLS results show that a positive productivity shock in the first layer of the hierarchy is associated with an increase in the productivity of the subsidiaries in the lower layers. These results are consistent with our model of knowledge hierarchies, where a BG's production efficiency percolates through the hierarchy of its subsidiaries from top to bottom.

[24] The data in the Historical Ownership Database of Bureau van Dijk starts in 2007.
[25] We measure labor productivity as $log(\text{operating revenue turnover}/\text{employees})$.

Table C1. Δ Bottom-layer productivity on Δ first-layer productivity.

Panel A: All groups

	***Dep. Var.*: Δ Prod 3 bottom layers**				
	(1)	(2)	(3)	(4)	(5)
Δ Prod First Layer	0.851***	0.856***	0.851***	0.775***	0.773***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)
FE		HQ	C	Y	C, Y
Observations	39,282	35,639	39,282	39,282	39,282

Panel B: Groups with at least four layers

	Dep. Var.: Δ Prod 3 bottom layers				
	(1)	(2)	(3)	(4)	(5)
Δ Prod First Layer	0.455***	0.499***	0.453***	0.132***	0.130***
	(0.021)	(0.024)	(0.021)	(0.025)	(0.025)
FE		HQ	C	Y	C, Y
Observations	5265	4491	5265	5265	5265

Notes: OLS estimations. In both Panels A and B, the dependent variable is the mean yearly growth in labor productivity of the three bottom layers, where labor productivity is measured as $log(\text{operatingrevenueturnover}/\text{employees})$. The independent variable (Δ Prod First Layer) is the mean yearly growth in labor productivity of the subsidiaries in the first layer. Standard errors are clustered at the BG's level. In both panels, column 1 does not include FEs, column 2 includes BG fixed effects, column 3 includes BG's country fixed effects, column 4 includes year fixed effects, and column 5 includes both country and year fixed effects. Panel A reports the estimates using all groups, while Panel B restricts the sample to BGs with at least four layers.
***p<.01.

APPENDIX D: MODEL SOLUTION

This appendix solves the model and provides a detailed derivation of the results described in Section 4.

D.1 Market structure and output choice

Consider an integrated global market in which a large number of BGs compete by selling imperfectly substitutable products. Each BG can produce only a countable number of products, which is a subset of measure zero of the total mass of products available in the market. This supports a monopolistic competition outcome even though BGs are multi-product suppliers (Mayer et al. 2014). For each of its products a BG faces isoelastic demand $y = Ap^{-\sigma}$, where y is output, p is price, $\sigma > 2$ is the cross-price as well as the own-price demand elasticity, and $A > 0$ is a demand shifter. A product's revenues are then given by $R(y) = py = A^{1/\sigma} y^{(\sigma-1)/\sigma}$.

Given marginal cost $1/\Omega$ and fixed cost W, a product's profit is

$$\Pi(y) = R(y) - y/\Omega R(y) - y/\Omega - W.$$

This is maximized when operating profits $R(y) - y/\Omega(\wp)$ are maximized. Profit maximizing output therefore satisfies the first order condition (FOC)

$$\Pi_y(y) = R_y(y) - 1/\Omega = 0,$$

where subscript y is used to refer to the derivatives with respect to output. Given the expression of $R(y)$, the FOC can be restated as

$$\frac{\sigma-1}{\sigma}A^{1/\sigma}y^{-\frac{1}{\sigma}} = 1/\Omega$$

so that optimal output evaluates to

$$y^* = A\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\Omega^{\sigma}$$

Then, maximized profit with respect to output is:

$$\begin{aligned}\Pi(y^*) &= R(y^*) - y^*/\Omega - W \\ &= a\Omega^{\sigma-1} - W\end{aligned}$$

where $a = (A/\sigma)[(\sigma-1)/\sigma]^{\sigma-1}$ is a bundle of demand parameters.

D.2 Model solution by backward induction

The optimal decisions that maximize the BG's profit can be characterized through backward induction. First, conditional on being led by an executive of type m who has solved a problem version $\wp$, the profit that a subsidiary earns from output y is revenues $R(y)$ net employees' wages $y/\Omega(\wp)$ and the executive's hiring cost $W(m)$. Due to $R(y) = A^{1/\sigma}y^{(\sigma-1)/\sigma}$, maximized profit with respect to output evaluates to $\Pi_1(\wp, m) = a\Omega(\wp)^{\sigma-1} - W(m)$. Second, conditional on the subsidiary having been assigned problem version $\wp$, the fact that $W(m)$ increases with the executive's ability and an executive of any ability $m \le \wp$ can find the solution implies that the hired executive has the minimum required ability $m = \wp$. Recalling $\Omega(\wp) = \omega e^{-\theta\wp}$ and $W(m) = we^{-\theta m}$, the profit the parent anticipates to earn by assigning problem version $\wp$ to the subsidiary can then be stated as

$$\Pi_1(\wp) = [a\omega^{\sigma-1}e^{-\theta(\sigma-2)\wp} - w]e^{-\theta\wp} \tag{D1}$$

With $\sigma > 2$ the demand elasticity is large enough for the additional profit from solving a harder problem version more than compensates the additional hiring cost for a more able executive who can solve it. As a result, $\Pi_1(\wp)$ is a decreasing function of $\wp$ as long as it positive. Third, given (D1), the parent maximizes the layer's aggregate profit

$$n_1\Pi_1(\wp) - wF = n_1[a\omega^{\sigma-1}e^{-\theta(\sigma-2)\wp} - w]e^{-\theta\wp} - wF \tag{D2}$$

with respect to the number of subsidiaries n_1 and the problem version $\wp$ subject to the advising time constraint $n_0\widehat{T} = n_1T(\wp, m)$. With $T(\wp, m) = e^{-\theta\wp}e^{\varphi m}$, the time constraint can be rewritten as

$$n_1 = e^{\tau+(\theta-\varphi)\wp}n_0$$

with $n_0 = 1$ as only the parent advises the first layer. Substituting the constraint into (D2) gives

$$n_1(\wp)\Pi_1(\wp) - wF = e^{\tau + (\theta - \varphi)\wp} n_0 [a\omega^{\sigma - 1} e^{-\theta(\sigma - 2)\wp} - w] e^{-\theta\wp} - wF \tag{D3}$$

Accordingly, $\sigma > 2$ is a sufficient condition for the parent to assign the most difficult problem ($\wp = 1$) to the first layer as it implies that $n_1(\wp)\Pi_1(\wp)$ is a decreasing function of $\wp$ as long as it is positive. However, for such assignment to make sense, it is also necessary that the first layer breaks at least even ($n_1(\wp)\Pi_1(\wp) - wF > 0$ for $\wp = 1$), which is the case when the parent-specific productivity component ω is large enough:

$$\omega > \omega_1 \equiv [\frac{w}{a}\left(1 + \frac{F}{e^{\tau - \varphi}}\right) e^{\theta(\sigma - 2)}]^{\frac{1}{\sigma - 1}}.$$

Under this condition, the profit maximizing choice with respect to the first layer entails $\ell = \wp = m = 1$ so that the first layer is equipped with executives that can advise the solution of any problem version $\wp > 1$.

Repeating the analysis for the second layer determines its profit as

$$n_2(\wp)\Pi_2(\wp) - wF = e^{\tau + (\theta - \varphi)\wp} n_1 [a\omega^{\sigma - 1} e^{-\theta(\sigma - 2)\wp} - w] e^{-\theta\wp} - wF$$

which for $\sigma > 2$ is also decreasing in $\wp$. However, the hardest problem version cannot be assign to this layer as executives of type $m = 1$ in the first layer 1 are not able enough to offer advice and the parent's time is completely absorbed by advising the first layer. Iterating the argument for the subsequent layers reaches an analogous conclusion: the third hardest problem version is assigned to the third layer, the fourth hardest problem version to the fourth layer, and so on.

To summarize, for $\sigma > 2$ and $\omega > \omega_1$, the layer index coincides with the problem version index, which in turn is matched by the executive type index ($\ell = \wp = m$). The number of subsidiaries at generic layer ℓ is then determined by the advising time constraint for layer $\ell - 1$, which equalizes, for $\ell = \wp = m$, the advising time needed by the former layer $n_\ell T(\ell, \ell) = n_\ell e^{(\varphi - \theta)\ell}$ and the latter layer's endowment $n_{\ell - 1}\widehat{T} = n_{\ell - 1} e^{\tau}$. The constraint thus implies that n_ℓ and $n_{\ell - 1}$ are linked by the relation $n_\ell = n_{\ell - 1} e^{\tau + (\theta - \varphi)}$. Applying this relation iteratively from $n_0 = 1$ up to n_ℓ yields the optimal number of subsidiaries

$$n_\ell = e^{\ell\tau + \frac{1}{2}\ell(\ell + 1)(\theta - \varphi)} \tag{D4}$$

for the generic layer ℓ, which is activated as long as it generates positive profits

$$n_\ell \Pi_\ell - wF = e^{\tau + (\theta - \varphi)\ell} n_{\ell - 1} [a\omega^{\sigma - 1} e^{-\theta(\sigma - 2)\ell} - w] e^{-\theta\ell} - wF > 0. \tag{D5}$$

Finally, the BG's hierarchy stops at some cutoff layer ℓ^* such that $n_\ell \Pi_\ell - wF \geq 0$ and $n_\ell \Pi_\ell - wF < 0$ simultaneously hold. The BG's total number of subsidiaries and its overall profits are then given by $N = \sum_{\ell=0}^{\ell^*} n_\ell$ and $\Pi = \sum_{\ell=0}^{\ell^*} (n_\ell \Pi_\ell - wF)$, respectively.

D.3 Comparative statics

Despite its parsimony, the model has rich comparative statics implications. For $\ell^* \leq P$, the cutoff layer determines the depth of the hierarchy with larger ℓ^* describing a deeper hierarchy with a larger number of layers. As a is an increasing function of ω, a BG's hierarchical depth increases with its production efficiency. Depth depends also on the rate at which layer profit

$n_\ell \Pi_\ell$ falls as ℓ rises, which is higher for lower advising efficiency (smaller τ) and lower communication efficiency (larger φ) as inferred from (D4) and (D5). Accordingly, a BG's hierarchical depth increases with its advising efficiency and communication efficiency. Hence we have:

Remark 1. ***(Number of layers)*** *The cutoff layer ℓ^* is an increasing function of ω and τ whereas it is a decreasing function of φ.*

Through (D4) advising efficiency and communication efficiency also affect the number of subsidiaries placed at each layer:

Remark 2. ***(Subsidiaries per layer)*** *The number of subsidiaries n_ℓ at layer ℓ is an increasing function of τ and a decreasing function of φ for all $\ell = 1, \ldots, \ell^*$.*

The same parameters determine the shape of the hierarchy. In particular, the model can generate pyramids (i.e., left-skewed hierarchies), inverse pyramids (i.e., right-skewed hierarchies) and mixed structures ('diamonds'). This is due to the fact that (D4) implies that a generic layer $\ell \geq 2$ has more (fewer) subsidiaries than the layer $\ell - 1$ above it if and only if $\tau + \ell(\theta - \varphi) > (<) 0$. Therefore, there exists a threshold value $\ell_d \equiv \tau/(\varphi - \theta)$ such that the hierarchy is a pyramid for $\ell < \ell_d$ and an inverse pyramid for $\ell > \ell_d$. This leads to:

Remark 3. ***(Shape)*** *For $\varphi < \theta$ the optimal hierarchical structure of a BG is a pyramid. For $\varphi > \theta$ the optimal shape of a BG's hierarchy is: a pyramid iff $\tau/(\varphi - \theta) \leq 2$, a diamond iff $2 < \tau/(\varphi - \theta) < \ell^*$, or an inverse pyramid iff $\tau/(\varphi - \theta) \geq \ell^*$.*

As a corollary, while a sufficient condition for a pyramid is $\theta > \varphi$, a sufficient condition for an inverse pyramid is $\tau + (\theta - \varphi) < 0$ given that it implies $\tau + \ell^*(\theta - \varphi) < 0$ for any $\ell^* \geq 1$. Remark 3 implies that, for given intermediate production efficiency gains (θ), a pyramid maximizes the overall profit of BGs with high advising efficiency (large τ) and high communication efficiency (small φ). In contrast, an inverse pyramid maximizes the overall profit of BGs with low advising efficiency (small τ) and low communication efficiency (large φ). A diamond is the best option for intermediate advising efficiency and communication efficiency. The larger these are, the larger the pyramidal part of the diamond.

Finally, (D4) sheds light on how steep pyramids or the pyramidal components of diamonds are. If we measure steepness by the absolute value of the difference between n_ℓ and $n_{\ell-1}$, then this is smaller the closer φ is to θ, leading to:

Remark 4. ***(Steepness)*** *The gap in the number of subsidiaries between layers $|n_\ell - n_{\ell-1}|$ is an increasing function of $|\theta - \varphi|$.*

We have seen that advising the solution of more difficult versions absorbs more time; however, as more difficult versions can be tackled only by higher ability advisees, less time is needed for communication. Then Proposition 4 states that pyramids or the pyramidal components of diamonds are steeper when those two opposing effects are close to offsetting each other (which happens for $\theta = \varphi$).

APPENDIX E: CONTRACTABILITY AND HIERARCHY

In Table E1 for each subsidiary we regress the contractability of its activities on its layer within its group hierarchy plus parent fixed effects. We use the contractability index developed by Nunn (2007). This index combines a measure of the quality of contract enforcement for each country with a measure of contractual intensity for each final good in the manufacturing sector (i.e., the share of inputs requiring relationship-specific investments).[26] Contractability varies by country and industry: a high value of contractability is associated with a high level of judicial quality combined with a high level of contract intensity. We define the contractability of a subsidiary's activities as the contractability index associated with its industry and its country of incorporation. Figure E1 plots the coefficients of the regression with their 95% confidence intervals and does not reveal any significant pattern of correlation between contractability and hierarchical distance from the parent.

Table E1. Contractability.

Dependent variable **Estimation method**	**Contractability** **OLS**
Layer 2	0.059***
	(0.002)
Layer 3	0.095***
	(0.004)
Layer 4	0.113***
	(0.005)
Layer 5	0.093***
	(0.009)
Layer 6	0.112***
	(0.012)
Layer 7	0.092***
	(0.017)
Layer 8	0.087***
	(0.025)
Layer 9	0.049
	(0.033)
Layer 10	0.101***
	(0.033)
Layer >10	0.150***
	(0.036)
Parent FE	Yes
Country FE	–
Observations	316,700
R^2	0.693

Notes: OLS estimations. *Contractability* is the contractability index developed by Nunn (2007). The index combines a measure of the quality of contract enforcement for each country with a measure of contractual intensity for each final good in the manufacturing sector. Contractability varies by country and sector: a high value of contractability is associated with a high level of judicial quality combined with a high level of contract intensity. *Layer* is the degree of separation between the HQ and the corresponding subsidiary. The constant is omitted from the table. All specifications include parent company FE. The model is estimated on the full world sample, excluding BGs with only one affiliate. Heteroskedasticity robust standard errors at the HQ level are in parentheses.
*** $p < .001$.

[26] Specifically, we measure contractability as $RoL \times z(a, b)$, where RoL is the 'Rule of Law' index from the Worldwide Governance Indicators (World Bank Group: Kaufmann et al. (2010)), and $z(a, b)$ is a variable from Nunn (2007), which measures the level of contract intensity of a specific industry; a represents the Rauch (1999) industry definition (liberal or conservative); b represents the methodology adopted to define relationship-specific inputs. The latter in turn can vary between inputs that are neither sold on an organized exchange nor reference priced, and those that are not sold on an organized exchange but are reference priced (see Nunn [2007] for details). Hence, the final measure of contractability encompasses four different indexes: the two measures of proportion of intermediate inputs and the two estimates by Rauch (1999).

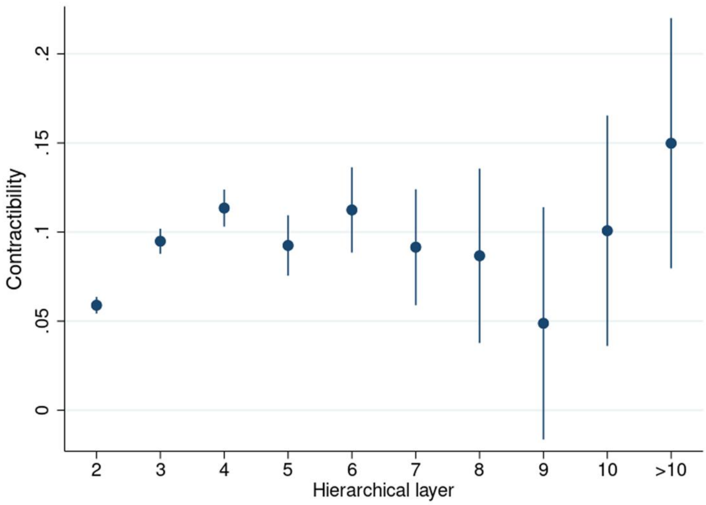

Figure E1. Contractability over hierarchies. The graph shows the coefficients and 95% confidence intervals obtained from a regression of the contractability index of each subsidiary (liberal definition, using inputs that are neither sold on an organized exchange nor reference priced) on the hierarchical layer of the same subsidiaries, including group FE, and using robust standard errors. Host-country FE are omitted because contractability incorporates rule of law that is country specific. Regression results are reported in Table E1.

ACKNOWLEDGMENTS

We are indebted to Luis Garicano for initial discussions and Davide Suverato for constant interactions. We are also grateful to Pol Antràs, Lorenzo Caliendo, Paola Conconi, Miguel Espinosa, Rocco Macchiavello, Esteban Rossi-Hansberg, Vincenzo Scrutinio, as well as seminar participants at the 2022 CEPR Workshop on Strategy and Structure of BGs, the 2018 ERWIT-CEPR meeting in St. Gallen, and the 2018 Princeton Summer Institute for useful comments. We thank Bianca Brunori, Paolo De Rosa, Giulia Lo Forte, Francesco Losma, Marta Mojoli, Alessandro Pisa, and Davide Zufacchi for excellent research assistance.

FUNDING

We acknowledge funding by the European Union's Horizon 2020 research and innovation program, grant agreement No. 822390 (MICROPROD).

Conflict of interest statement. None declared.

REFERENCES

Alfaro, L. 2017. "Gains from Foreign Direct Investment: Macro and Micro Approaches," 30 *The World Bank Economic Review* S2–S15.

Alfaro, L., and A. Charlton. 2009. "Intra-Industry Foreign Direct Investment," 99 *American Economic Review* 2096–119.

Almeida, H., S. Y. Park, M. G. Subrahmanyam, and D. Wolfenzon. 2011. "The Structure and Formation of Business Groups: Evidence from Korean Chaebols," 99 *Journal of Financial Economics* 447–75.

Almeida, H. V., and D. Wolfenzon. 2006. “A Theory of Pyramidal Ownership and Family Business Groups,” 61 *The Journal of Finance* 2637–80.

Altomonte, C., and A. Rungi. 2013. “Business Groups as Hierarchies of Firms: Determinants of Vertical Integration and Performance.” Technical report, ECB Working Paper. European Central Bank.

Aminadav, G., and E. Papaioannou. 2020. “Corporate Control around the World,” 75 *The Journal of Finance* 1191–246.

Bartlett, C. A., and S. Ghoshal. 1989. *Managing across Borders: The Transnational Solution*. Brighton (MA): Harvard Business School Press.

Belenzon, S., and T. Berkovitz. 2010. “Innovation in Business Groups,” 56 *Management Science* 519–35.

Belenzon, S., N. Hashai, and A. Patacconi. 2019. “The Architecture of Attention: Group Structure and Subsidiary Autonomy,” 40 *Strategic Management Journal* 1610–43.

Belenzon, S., and M. Schankerman. 2013. “Spreading the Word: Geography, Policy, and Knowledge Spillovers,” 95 *The Review of Economics and Statistics* 884–903.

Björkman, I., W. Barner-Rasmussen, and L. Li. 2004. “Managing Knowledge Transfer in MNCs: The Impact of Headquarters Control Mechanisms,” 35 *Journal of international business studies* 443–55.

Blinder, A. S., and A. B. Krueger. 2009, August. Alternative measures of offshorability: A survey approach. Working Paper 15287, National Bureau of Economic Research.

Blinder, A. S., and A. B. Krueger. 2013. “Alternative Measures of Offshorability: A Survey Approach,” 31 *Journal of Labor Economics* S97–128.

Bresman, H., J. Birkinshaw, and R. Nobel. 1999. “Knowledge Transfer in International Acquisitions,” 30 *Journal of International Business Studies* 439–62.

Cestone, G., C. Fumagalli, F. Kramarz, and G. Pica. 2023. “Exploiting Growth Opportunities: The Role of Internal Labour Markets,” 91 *The Review of Economic Studies* 2676–2716.

Chapelle, A., and A. Szafarz. 2007. “Control Consolidation with a Threshold: An Algorithm,” 18 *IMA Journal of Management Mathematics* 235–43.

Ciabuschi, F., H. Dellestrand, and P. Kappen. 2011. “Exploring the Effects of Vertical and Lateral Mechanisms in International Knowledge Transfer Projects,” 51 *Management International Review* 129–55.

Cravino, J., and A. A. Levchenko. 2017, November. “The Distributional Consequences of Large Devaluations,” 107 *American Economic Review* 3477–509.

Crespo, C. F., L. F. Lages, and N. F. Crespo. 2020. “Improving Subsidiaries’ Innovation through Knowledge Inflows from Headquarters and Peer Subsidiaries,” 26 *Journal of International Management* 100803.

Del Prete, D., and A. Rungi. 2017. “Organizing the Global Value Chain: A Firm-Level Test,” 109 *Journal of International Economics* 16–30.

Eppinger, P. S., and B. Kukharskyy. 2020. “Contracting Institutions and firm Integration around the World.” Technical report, University of Tübingen Working Papers in Economics and Finance.

EUROSTAT. 2007. Recommendations Manual on the Production of Foreign Affiliates Statistics (FATS). Publications Office of the European Union.

Faccio, M., and L. H. Lang. 2002. “The Ultimate Ownership of Western European Corporations,” 65 *Journal of financial economics* 365–95.

Fadeev, E. 2023. “Creative Construction: Knowledge Sharing and Cooperation Between Firms.” Technical report, Working paper, Duke University Fuqua School of Business.

Garg, G., M. Sewak, and A. Sharma. 2022. “Learning from Older Siblings: Impact on Subsidiary Performance,” 31 *International Business Review* 101957.

Garicano, L., and E. Rossi-Hansberg. 2015. “Knowledge-Based Hierarchies: Using Organizations to Understand the Economy,” 7 *Annual Review of Economics* 1–30.

Gorodnichenko, Y., B. Kukharskyy, and G. Roland. 2021. Cultural Distance, Firm Boundaries, and Global Sourcing. Technical report, UC Berkeley.

Grosskurth, P. 2019. “MNE and Where to Find Them: An Intertemporal Perspective on the Global Ownership Network.” Technical report, Ruhr Economic Papers.

Gupta, A. K., and V. Govindarajan. 1991. “Knowledge Flows and the Structure of Control within Multinational Corporations,” 16 *Academy of management review* 768–92.

Huneeus, F., B. Larrain, M. Larrain, and M. Prem. 2021. “The Internal Labor Markets of Business Groups,” 69 *Journal of Corporate Finance* 102017.

IFRS. 2011. International Financial Reporting Standard N.10 on Consolidated Financial Statements.

Khanna, T., and K. Palepu. 2000. "Is Group Affiliation Profitable in Emerging Markets? An Analysis of Diversified Indian Business Groups," 55 *The Journal of Finance* 867–91.
Khanna, T., and J. W. Rivkin. 2001. "Estimating the Performance Effects of Business Groups in Emerging Markets," 22 *Strategic Management Journal* 45–74.
La Porta, R., F. Lopez-de Silanes, and A. Shleifer. 1999. "Corporate Ownership around the World," 54 *The Journal of Finance* 471–517.
Lewellen, K., and L. A. Robinson. 2013. "Internal Ownership Structures of US Multinational Firms," working paper, available at ssrn.com.
Li, J., and R. P. Lee. 2015. "Can Knowledge Transfer within MNCs Hurt Subsidiary Performance? The Role of Subsidiary Entrepreneurial Culture and Capabilities," 50 *Journal of World Business* 663–73.
Mayer, T., M. J. Melitz, and G. I. P. Ottaviano. 2014. "Market Size, Competition, and the Product Mix of Exporters," 104 *American Economic Review* 495–536.
Monteiro, L. F., N. Arvidsson, and J. Birkinshaw. 2008. "Knowledge Flows within Multinational Corporations: Explaining Subsidiary Isolation and Its Performance Implications," 19 *Organization Science* 90–107.
Nair, S. R., M. Demirbag, and K. Mellahi. 2015. "Reverse Knowledge Transfer from Overseas Acquisitions: A Survey of Indian MNEs," 55 *Management International Review* 277–301.
Nunn, N. 2007. "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," 122 *The Quarterly Journal of Economics* 569–600.
OECD. 2005. *Guidelines for Multinational Enterprises*. Paris: OECD Publishing.
OECD. 2008. *OECD Benchmark Definition of Foreign Direct Investment*, 4th edition. Paris: OECD Publishing.
Persson, M. 2006. "The Impact of Operational Structure, Lateral Integrative Mechanisms and Control Mechanisms on intra-MNE Knowledge Transfer," 15 *International Business Review* 547–69.
Rauch, J. E. 1999. "Networks versus Markets in International Trade," 48 *Journal of international Economics* 7–35.
Rungi, A., G. Morrison, and F. Pammolli. 2017. "Global Ownership and Corporate Control Networks," 7/2017 *IMT Lucca EIC WP Series* 1–34.
Sonno, T. 2025. "Globalization and Conflicts: The Good, the Bad, and the Ugly of Corporations in Africa," 135 *The Economic Journal* 1108–40.
Subramaniam, M., and N. Venkatraman. 2001. "Determinants of Transnational New Product Development Capability: Testing the Influence of Transferring and Deploying Tacit Overseas Knowledge," 22 *Strategic Management Journal* 359–78.
UNCTAD. 2009. The UNCTAD Manual on Statistics of FDI and the Operations of TNCs. United Nations, Geneva.
UNCTAD. 2016. World Investment Report 2016 - Investor Nationality: Policy Challenges. United Nations.
Wellman, N., J. M. Applegate, J. Harlow, and E. W. Johnston. 2020. "Beyond the Pyramid: Alternative Formal Hierarchical Structures and Team Performance," 63 *Academy of Management Journal* 997–1027.
World Bank Group: D., Kaufmann, A. Kraay, and M. Mastruzzi. 2010. Worldwide Governance Indicators. World Bank Group.
Wu, Y., R. Strange, and V. Shirodkar. 2022. "Lateral Knowledge Transfer and Foreign Affiliate Performance: The Importance of Affiliate Strategic Roles," 116 *Economic Modelling* 106039.

The Journal of Law, Economics, and Organization, 2026, 00, 1–50
https://doi.org/10.1093/jleo/ewag011
Article