

Jahn, Elke; Egger, Hartmut; Kornitzky, Stefan

Conference Paper

Does the Position in Business Group Hierarchies Affect Workers' Wages?

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Jahn, Elke; Egger, Hartmut; Kornitzky, Stefan (2021) : Does the Position in Business Group Hierarchies Affect Workers' Wages?, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242374>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Does the Position in Business Group Hierarchies Affect Workers' Wages?*

Hartmut Egger[§]
*University of Bayreuth
CESifo, GEP and IfW*

Elke Jahn
*IAB Nuremberg, University
of Bayreuth and IZA*

Stefan Kornitzky
University of Bayreuth

February 16, 2021

Abstract

We merge firm-level data on ownership linkages with administrative data on German workers to analyze how the position in a business group hierarchy affects workers' wages. To acknowledge that ownership linkages are not one-directional, we propose an index to measure hierarchical distance to the ultimate owner. After controlling for unobserved heterogeneity and selection into the business group hierarchy, we find a positive effect of larger hierarchical distance to the ultimate owner of a business group on workers' wages. This result is in line with a monitoring-based theory of business group hierarchies, in which larger hierarchical distance to the ultimate owner is associated with lower monitoring efficiency and thus higher wage payments to incentivize workers.

JEL-Classifications: C23, J31, L23

Keywords: Business groups, ownership networks, workers wages, difference-in-difference, hierarchical distance

*We would like to thank participants of the Graduate Seminar at the University of Bayreuth and the Bavarian Graduate Program in Economics Workshop for valuable discussion.

[§]Corresponding author: University of Bayreuth, Department of Economics, Universitaetsstr. 30, 95447 Bayreuth, Germany; hartmut.egger@uni-bayreuth.de.

1 Introduction

Although it is widely perceived that business groups account for a major part of economic activity (UNCTAD, 2016), as a hybrid form between firms and markets they are not well defined objects of economic theory and have therefore so far received little attention in academic research (see Baker et al., 2002; Khanna and Yafeh, 2007; Altomonte et al., 2018). In particular, it is not understood yet, how the specific organizational form of a business group influences its economic performance and workers' wages. This is surprising given the vast empirical evidence showing that firm organization is a key determinant of productivity and wages (see Caroli and Van Reenen, 2001; Rajan and Wulf, 2006; Bloom et al., 2010, 2018). Based on recent empirical findings that the vertical position of workers in firm hierarchy plays a crucial role for their wages (see Caliendo et al., 2015; Bastos et al., 2018; Friedrich, forthcoming), we ask whether a similar effect on workers' wages can be found for the position of their employer in a business group hierarchy.¹

For our analysis, we define business groups as ownership networks and assume that the ultimate owner exercises hierarchical control over the decisions made in all affiliated firms. We extract the relevant ownership information from the Bureau van Dijk global firm database Orbis. This database provides insights on worldwide ownership linkages and thus gives detailed information on the hierarchical position of firms in their business groups. To determine how the hierarchical position impacts workers' wages, we merge Orbis with administrative data on German employees from the Institute for Employment Research (IAB) in Nuremberg. This gives us a novel dataset with detailed information on the business group German establishments are part of and rich data on their workforce.² As a defining feature of our dataset, we observe that ownership linkages are not one-directional, so that a simple count of hierarchical layers between an establishment and its ultimate owner would give at best an imprecise measure of vertical distance. To capture the complex structure of ownership networks in our data, we develop a hierarchical distance index, which is motivated by recent work on sectoral input-output relationships (cf. Antràs and

¹Business groups play a prominent role in a sizable, mostly empirical literature on foreign ownership wage premia (see Girma et al., 2001; Girma and Görg, 2007; Balsvik and Haller, 2010; Hijzen et al., 2013; Egger et al., 2020; Egger and Jahn, 2020). Whereas this literature emphasizes the geographical location of the ultimate owner as an important determinant of wages in foreign subsidiaries, we study how the hierarchical position in business groups affects workers' wages.

²In the administrative data of Germany workers are employed by establishments, whereas in Orbis the observational units are firms, which depending on their hierarchical position in the business group can be further divided into ultimate owners and subsidiaries. We match establishment-level administrative data on workers with firm-level observations from Orbis using record linkages.

Chor, 2013) and measures hierarchical distance more consistently than a pure count of ownership layers.

To obtain intuition on how the position in business groups affects workers' wages, we set up a theoretical model, in which production requires consecutive performance of a continuum of stages along the value chain of the business group. The value chain is split into two segments of endogenous length, which are operated by an upstream and a downstream firm (cf. Costinot et al., 2012). Crucial for our analysis, we assume that the production process is prone to a loss of control problem due to limited monitoring capacity of the ultimate owner (see Calvo and Wellisz, 1979; Chen, 2017). Focussing on the problem of a single business group, we show that optimal labor allocation and wage profile depend on relative monitoring efficiency in the upstream and downstream firm. Assuming that the value chain follows the hierarchical structure of the business group, which is common practice, for instance, in the context of vertical multinational enterprises, we associate the upstream producer with a lower layer in business group hierarchy. In this case, lower monitoring efficiency in the hierarchically more distant upstream firm leads to a positive impact of hierarchical distance on wages, whereas the opposite is true if monitoring efficiency is lower in the hierarchically less distant downstream firm.

3

In the empirical analysis, we control for observable worker and establishment characteristics, such as age, gender and education, but also establishment size and industry affiliation to isolate the effects of hierarchical distance from other factors that have shown to be important for wage payments by previous empirical research. Including these controls, we identify a positive marginal effect of larger hierarchical distance to the ultimate owner of a business group on individual wages in German establishments. According to our estimates from a parsimonious OLS specification,

³Our model shows close resemblance with the monitoring-based theory of firm hierarchies that has been prominently discussed by Calvo and Wellisz (1978, 1979), Qian (1994), and Chen (2017). According to this theory, hierarchical layers can alleviate the loss of control problem by increasing monitoring capacity and thereby reducing the incentive pay necessary to align workforce behavior with the objective of the owner. An assumption common to existing models of monitoring-based hierarchies is that production is concentrated at the lowest hierarchical layer, while workers employed at higher layers exclusively perform monitoring tasks. This feature is at odds with empirical evidence on business groups (see Hanson et al., 2005; Atalay et al., 2014) and thus eliminated in our model. Our model is also related to the knowledge-based theory of firm hierarchies, in which hierarchical layers facilitate the information flow between workers and their superiors and thereby reduce the number of unsolved problems in the production process (see Garicano, 2000; Garicano and Rossi-Hansberg, 2006; Caliendo and Rossi-Hansberg, 2012). Chen and Suen (2019) discuss differences and similarities between monitoring-based and knowledge-based theories of firm hierarchies.

an increase in the hierarchical distance by one standard deviation amounts to a sizable increase in wages of almost two log points. Although this estimate is reduced when additionally controlling for unobserved worker, establishment, and business group heterogeneity by fixed-effects, a positive and significant effect of hierarchical distance on wages still exists under this alternative specification.

To make sure that the hierarchical distance variable does not erroneously pick up other features of business groups, we control for the total number of subsidiaries, as suggested by rich evidence for a firm size-wage premium (see, for instance, Brown and Medoff, 1989; Idson and Oi, 1999; Winter-Ebmer and Zweimüller, 1999; Colonelli et al., 2018). In addition, we combine information on the horizontal and the vertical dimension of business groups to an entropy index, measuring business group complexity (see Altomonte and Rungi, 2015). Adding this covariate, the impact of hierarchical distance on wages remains positive with OLS as well as fixed-effects regressions. Splitting the sample into three skill groups, we find a positive and significant effect of larger hierarchical distance on the wage of low- and medium-skilled workers, whereas the evidence for high-skilled workers is less clear. We also address the potential problem of selection bias by combining propensity-score matching with a difference-in-difference estimator. This two-stage procedure gives a picture that is broadly in line with our estimation results from OLS and fixed-effects regressions. Larger hierarchical distance increases wages on average as well as for the individual skill groups. Against the background of our theoretical model, the empirical results indicate that larger hierarchical distance is associated with lower monitoring efficiency and thus with higher wages necessary for incentivizing workers to follow the objectives of the ultimate owner of the business group.

The remainder of this paper is organized as follows. In Section 2, we outline a theoretical model for explaining wage payments along the business group hierarchy. In Section 3, we explain how we merge global firm data from Orbis with administrative data of German workers from the IAB. There, we also report summary statistics and show descriptive evidence on the relationship between hierarchical distance to the ultimate owner of the business group and workers' wages. In Section 4, we present the empirical analysis and report our estimation results, Section 5 concludes.

2 A theoretical account

Let us consider a single business group in a competitive market that can sell its output at a given price equal to one. The business group operates a continuum of consecutively performed production stages with measure one (see Costinot et al., 2012) and faces the trade-off between monitoring workers or paying higher wages to reduce shirking (cf. Calvo and Wellisz, 1979; Chen, 2017). To facilitate our analysis, we consider a simple structure with two firms, which are associated with an upstream (intermediate goods) producer, $j = u$, and a downstream (final goods) producer, $j = d$, respectively. The value chain of the business group is split between these two firms into two disjoint segments with endogenous length. Capturing the value chain by the unit interval, we denote by $S \in (0, 1)$ the segment performed by the upstream producer and by $1 - S$ the segment performed by downstream producer.⁴ The ultimate owner of the business group makes all relevant decisions on production, hiring, and monitoring for both firms.

Production technology

Following Costinot et al. (2012), we consider a Leontief technology that combines one unit of labor input with one unit of intermediate good from the previous stage to produce intermediate output. Thereby, $\lambda \in (0, 1)$ captures a Poisson rate at which mistakes occur and destroy output in the production of the two firms. For an infinitesimal ds , we can express the technology of producing stage $s + ds$ as

$$q(s + ds) = (1 - \lambda ds)q(s). \quad (1)$$

In the limit of $ds \rightarrow 0$, Eq. (1) establishes the differential equation $q'(s) = -\lambda q(s)$, whose solution is given by $q(s) = q(0) \exp(-\lambda s)$ and determines business group output at stage s as a function of the initial input $q(0)$, which we associate with a cost-free intangible asset of the business group.

We denote the accumulated production cost for one unit of output at stage s in firm j by $c_j(s)$. Accordingly, for an infinitesimal ds the costs of producing $q(s + ds)$ in firm j can be expressed as $c_j(s)q(s) + w_j q(s) ds$. Substituting $q(s + ds) = (1 - \lambda ds)q(s)$ from Eq. (1) gives $c_j(s + ds) = [c_j(s) + w_j ds]/(1 - \lambda ds)$, which in the limit can be expressed as the differential equation $c'_j(s) = \lambda c_j(s) + w_j$. Solving this differential

⁴Since in the limit the ultimate owner of the business group can assign a segment of length zero to either firm, the number of layers is endogenous in our model, despite its simple structure.

equation for either firm and making use of the boundary conditions $c_u(0) = 0$ and $c_u(S) = c_d(S)$, we can compute the labor costs of producing one unit of final output of the business group at $s = 1$ according to

$$c_d(1) = -\frac{w_d}{\lambda} + \left\{ \frac{w_u}{\lambda} [\exp(\lambda S) - 1] + \frac{w_d}{\lambda} \right\} \exp[\lambda(1 - S)] \equiv c.$$

Due to the Leontief technology, we can determine labor demand of the upstream and the downstream firm according to $\ell_u = \int_0^S q(s)ds$ and $\ell_d = \int_S^1 q(s)ds$, respectively. Solving these two integrals gives

$$S = -\frac{1}{\lambda} \ln \left[\frac{q(0) - \lambda \ell_u}{q(0)} \right] \quad \text{and} \quad \lambda(\ell_u + \ell_d) = q(0) \{1 - \exp(-\lambda)\}, \quad (2)$$

where $q(0)$ is intangible asset input, $q(0) - \lambda \ell_u$ is intermediate output of the upstream producer and $q(0) - \lambda(\ell_u + \ell_d) = q(0) \exp(-\lambda)$ is final output of the downstream producer. In view of Eq. (2), the unit cost of production simplifies to

$$c = \frac{w_u \ell_u + w_d \ell_d}{q(0) \exp(-\lambda)}. \quad (3)$$

Hiring, monitoring, and incentive pay

Firms hire workers at a convex cost of $\zeta \ell_j^2$.⁵ Workers have a disutility of effort equal to one and thus an incentive to shirk, which would lower their labor productivity to zero. The probability of a shirker to be detected by the ultimate owner of the business group (resulting in immediate job loss and zero income) is firm specific and given by $p_j = \alpha_j m_j / \ell_j$, where m_j is monitoring input while $\alpha_j > 0$ captures monitoring efficiency. The participation constraint of workers can be written as $w_j \geq 1/p_j = \ell_j / (\alpha_j m_j)$ and it holds with equality if the ultimate owner chooses the profit-maximizing wage. Similar to other models featuring a loss of control problem, we assume that the monitoring capacity of the ultimate owner of the business group is limited and normalized to one: $m_u + m_d = 1$ (see Calvo and Wellisz, 1979; Chen, 2017).

⁵Whereas it is important for our analysis that hiring costs can differ between the upstream and downstream firm, a quadratic form is not necessary and imposed for the sake of analytical tractability.

The optimization problem

We can study the business group's optimization problem in two steps. In step one, we solve for cost-minimizing labor and monitoring inputs, ℓ_j, m_j , holding output $q(0) \exp(-\lambda)$ constant. In step two, we then determine the profit-maximizing level of output, $q(0) \exp(-\lambda)$, given the business group's cost function.

Making use of the binding participation constraint $w_j = \ell_j/(\alpha_j m_j)$ and the production technology in Eq. (2), total (production plus hiring) costs can be expressed as

$$C(m_u, \ell_u, q(0)) \equiv \zeta \left\{ \left(\frac{a_u}{m_u} + 1 \right) \ell_u^2 + \left(\frac{a_d}{1 - m_u} + 1 \right) \left[\frac{q(0)}{\lambda} \{1 - \exp(-\lambda)\} - \ell_u \right]^2 \right\}.$$

where $a_j \equiv (\alpha_j \zeta)^{-1}$ is an auxiliary variable, inversely related to monitoring efficiency. Minimizing $C(m_u, \ell_u, q(0))$ for a given level of $q(0)$ establishes

$$\ell_u = \frac{a_d m_u + m_u(1 - m_u)}{a_u(1 - m_u) + a_d m_u + 2m_u(1 - m_u)} \frac{q(0)}{\lambda} \{1 - \exp(-\lambda)\} \quad (4)$$

and

$$m_u = \frac{\sqrt{a_u}(1 + a_d) - \sqrt{a_d a_u}}{\sqrt{a_u} + \sqrt{a_d}}. \quad (5)$$

Thereby, an interior solution with $m_u \in (0, 1)$ requires $\sqrt{a_u a_d} < 1 + \min\{a_u, a_d\}$ and thus the difference between a_u and a_d to be not too large. Using the solution to the cost-minimization problem, we can express total profits of the business group as

$$\Pi = q_o \exp(-\lambda) - \zeta \left(\frac{q(0)}{\lambda} \right)^2 \{1 - \exp(-\lambda)\}^2 \frac{(a_d + 1 - m_u^*)(a_u + m_u^*)}{a_u(1 - m_u^*) + a_d m_u^* + 2m_u^*(1 - m_u^*)},$$

where an asterisk is used to indicate the solution to the cost-minimization problem. Maximizing profits over $q(0)$ then gives an interior solution with $p_j < 1$ if ζ is sufficiently small.

Hierarchical wage profile

Differences in exogenous monitoring efficiency lead to differences in endogenous monitoring, according to Eq. (5). Moreover, it follows from Eqs. (2), (4), and (5) that higher monitoring efficiency is associated with higher labor input. We have $\ell_u > \ell_d$ if $\alpha_u > \alpha_d$, $\ell_u < \ell_d$ if $\alpha_u < \alpha_d$, and $\ell_u = \ell_d$ in the symmetric case of $\alpha_u = \alpha_d$. The

effect of monitoring efficiency on labor allocation follows from its effect on wages, which can be determined when noting that the optimal allocation of monitoring input is characterized by the condition

$$\frac{\ell_u}{\alpha_u m_u} = \sqrt{\frac{\alpha_d}{\alpha_u}} \frac{\ell_d}{\alpha_d (1 - m_u)}. \quad (6)$$

Making use of the binding participation constraint, Eq. (6) can be reformulated to $w_u = w_d \sqrt{\alpha_d / \alpha_u}$.⁶

Imposing the assumption that the position of a firm in the value chain is decisive for its position in business group hierarchy, which is common practice, for instance, in the literature on vertical multinational enterprises, we conclude that wages decrease along the value chain, i.e. $w_u > w_d$, if larger hierarchical distance to the ultimate owner of the business group is associated with lower monitoring efficiency, i.e. $\alpha_d > \alpha_u$. Existing research on organization networks gives good reason to believe that larger (hierarchical) distance is associated with higher costs of supervision (see Gumpert, 2018) – with the cost-saving motive providing a plausible explanation for the observed flattening of firm hierarchies over recent years (see Rajan and Wulf, 2006). However, the literature on empowerment supports the opposing view that larger hierarchical distance may reflect less commitment of workers, making monitoring more efficient (see Hill et al., 2012; Triguero-Sánchez et al., 2013). Consequently, larger hierarchical distance to the ultimate owner of the business group is associated with higher monitoring efficiency and thus lower wages. Since both arguments appear justified from a theory point of view, the impact of hierarchical distance on wages in business groups remains an empirical question that will be addressed below.

3 Data source and descriptives

We follow Altomonte and Rungi (2013) and associate business groups with ownership networks of legally autonomous firms. Moreover, we restrict attention to controlling ownership and therefore only consider ownership linkages above a 25 percent threshold. In cases, in which all shareholders hold a minority of ownership,

⁶Although we have assumed in our theoretical analysis that firms have identical labor productivity, the fundamental condition in Eq. (6) governing the hierarchical wage profile in the business group would remain unchanged if productivity differences existed. Productivity differences would affect, however, the labor allocation within the business group, with production increased, *ceteris paribus*, in the firm showing the lower Poisson rate of mistake, λ .

the firm with the highest minority share above 25 percent is assumed to exert control. Ownership captures upstream and downstream linkages and thus the vertical position of a firm in the business group hierarchy. In the following three subsections we introduce and describe our dataset. Thereby, we first explain how we combine information on business groups, firms, and workers from two different sources. We then introduce the main variables and provide summary statistics for them. In a final step, we show descriptive evidence for the link between the hierarchical distance to the ultimate owner of a business group and workers' wages.

3.1 Construction of the dataset

For our empirical analysis, we rely on two datasets. The first one covers the years 2013-2017 of Bureau van Dijk's commercial firm database Orbis. Orbis reports balance sheet information for several 100 million companies and their ownership linkages worldwide. Orbis covers all firms that are subject to reporting obligations. For Germany, these are all corporate enterprises and cooperatives as well as large private companies with total assets or revenues above thresholds defined by law. We select for each observation year German firms from Orbis that fulfill some minimum quality criteria and determine their ultimate owner, who can be German or not.⁷ We then extract the whole business group of the ultimate owner, using the available ownership information. For each firm within a business group we can determine the hierarchical layer, that is the number of (vertical) firm-linkages to the ultimate owner.

This procedure allows us to identify for each year about 40,000 different business groups, which cover at least one firm in Germany and represent in total almost a million firms worldwide.

As a second dataset, we use the Integrated Employment Biographies (IEB) from the Institute for Employment Research (IAB) of the German Federal Employment Agency. This dataset contains administrative records on all employees who are subject to social security contributions and covers about 80 percent of the German workforce. The IEB provides detailed information about age, gender, nationality, occupation, education, and the daily wage of workers employed in German establishments (see Klosterhuber et al., 2016). The IEB does not provide exact infor-

⁷Firms must be active and their legal form as well as their independence indicator have to be known. Moreover, operating revenues and number of employees have to be available for at least one year between 2012 and 2017. Finally, to build the relevant business group, we only keep firms with valid information on a unique ultimate owner.

mation on hours worked. Moreover, since worker information comes from social security records, wages are top-coded at the social security contribution ceiling. To deal with these issues, we consider only full-time workers aged 16–65 years for our analysis and impute wages above the social security contribution ceiling, using the two-step Tobit procedure provided by Dauth and Eppelsheimer (2020) – see Card et al. (2013) and Dustmann et al. (2014) for a similar approach.

To merge information on administrative data of German workers from IEB with firm-level information on business groups from Orbis, we employ the linkage procedure of the IAB, which makes use of company names and legal forms and is explained in detail by Antoni et al. (2018). To make sure that we correctly allocate establishments to firms over the whole sample period, we link establishments to firms separately for each year between 2013 and 2017. The resulting record linkage keys allow us to link on average 50,000 firms belonging to one of the business groups in Orbis with 86,000 establishments in IEB per year. Finally, to ensure that each full-time worker is uniquely linked to an establishment, we only keep employment spells that are valid on the 31st of December of a given year.

Firm-level variables from Orbis

Table 1 summarizes descriptive statistics on the main firm-level variables in our final dataset, which by construction do not vary over establishments or workers. For the whole sample period we count 250,494 firm-year observations.

The first two variables reported in Table 1 capture characteristics of the entire business group and are therefore identical for all firms belonging to the same group. The first variable is the total number of subsidiaries of a business group and therefore refers to *group size*. For the smallest business group we count only one subsidiary, whereas for the largest one, we count more than 13,000 subsidiaries. The average group size is 93 and thus fairly large. The second variable combines information on the number of subsidiaries and the number of ownership layers to an entropy index, which we refer to as *group complexity*(GC). It is constructed following Altomonte and Rungi (2015): $GC \equiv \sum_{l=1}^L l \frac{n_l}{N-1} \ln \left(\frac{N-1}{n_l} \right)$, where $N - 1$ is the total number of subsidiaries, L is the total number of ownership layers, and n_l is the number of subsidiaries at ownership layer $l \in \{1, \dots, L\}$. Group complexity picks up how the number of subsidiaries are spread over different ownership layers of a business group. It increases in the number of layers and places a higher weight on hierarchically more distant subsidiaries. Group complexity takes a minimum value of zero for business

groups with only one layer and is unbounded from above. Its maximum level in our dataset is 28.4.

Table 1: Business group characteristics and hierarchical distance

	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
Group size (in 100)	0.933	3.896	0.01	134.340
Group complexity	1.376	2.243	0	28.448
Hierarchical distance	1.245	1.178	0	18.770

Notes: Business group characteristics are constructed for the years 2013-2017, using firm-level information on ownership linkages from Bureau van Dijk's Orbis database. Group size is given by the total count of subsidiaries of a business group. Group complexity (GC) is defined following Altomonte and Rungi (2015). Data moments are reported for 250,494 firm-year observations.

The main control in our analysis is the hierarchical distance of a firm to its ultimate owner. To construct a sensible measure of hierarchical distance, we have to acknowledge that more than nine percent of the firm-year observations in our dataset show ownership linkages that are not one-directional. For instance, ownership linkages are bi-directional if firm j is partly owned by another firm in the business group that is in turn partly owned by firm j . However, ownership linkages can be even more complicated than that, because they often span over multiple layers of hierarchy, including many different firms. We account for this complex pattern by developing an index of hierarchical distance that captures circular network structures in a comprehensive way. To construct our index, we built on an index that has recently been applied for determining the vertical position of industries in global value chains (cf. Antràs et al., 2012; Antràs and Chor, 2013).

As a point of departure, we use the available ownership information, denote by ρ_{jk} the share of firm j that is owned by firm k , and express the chain of ownership in a business group as follows:

$$\sum_{k=1}^N \rho_{jk} + \sum_{k=1}^N \sum_{h=1}^N \rho_{jh} \rho_{hk} + \sum_{k=1}^N \sum_{h=1}^N \sum_{l=1}^N \rho_{jl} \rho_{lh} \rho_{hk} + \dots, \quad (7)$$

where N is the total number of firms in a business group, including the ultimate owner and all its subsidiaries. Thus, ρ_{jj} is zero by assumption. The first element of the series in (7) measures how the *direct* ownership of firm j is spread in the business group and therefore refers to the first level of outside control. The other elements refer to *indirect* ownerships, taking into account that firms holding shares of subsidiary j can in turn be owned by other firms in the business group. The series

in (7) is in general not bounded from above, $\sum_{k=1}^N \rho_{jk} > 1$. Therefore, we replace ρ_{jk} by $\hat{\rho}_{jk} \equiv \rho_{jk} / \sum_{k=1}^N \rho_{jk}$, and define hierarchical distance of firm j to its ultimate owner according to

$$H_j = \sum_{k=1}^N \hat{\rho}_{jk} + \sum_{k=1}^N \sum_{h=1}^N \hat{\rho}_{jh} \hat{\rho}_{hk} + \sum_{k=1}^N \sum_{h=1}^N \sum_{l=1}^N \hat{\rho}_{jl} \hat{\rho}_{lh} \hat{\rho}_{hk} + \dots$$

Using matrix notation, we can then summarise the hierarchical distance of all firms in a business group to their common ultimate owner by a single vector:

$$\mathbf{H} = \mathbf{R} \cdot \mathbf{1} + \mathbf{R}^2 \cdot \mathbf{1} + \mathbf{R}^3 \cdot \mathbf{1} \dots = [\mathbf{I} - \mathbf{R}]^{-1} \mathbf{1} - \mathbf{1}, \quad (8)$$

where $\mathbf{1}$ is an $N \times 1$ column vector of ones and \mathbf{R} is an $N \times N$ matrix with $\hat{\rho}_{jk}$ as its (j, k) -th element.⁸ The hierarchical distance of j from its ultimate owner is then given by the j -th row of the $N \times 1$ column vector \mathbf{H} . Eq. (8) determines the hierarchical distance of a subsidiary as a value-weighted count of the number of ownership layers between j and the ultimate owner in the business group (see Johnson, 2018). Higher values of H_j refer to a longer hierarchical distance and the index is normalized to give the ultimate owner a hierarchical distance value of zero. In our dataset, the hierarchical distance has a maximum of 18.8. If all ownership linkages were one-directional, index H_j would coincide with a simple layer count.

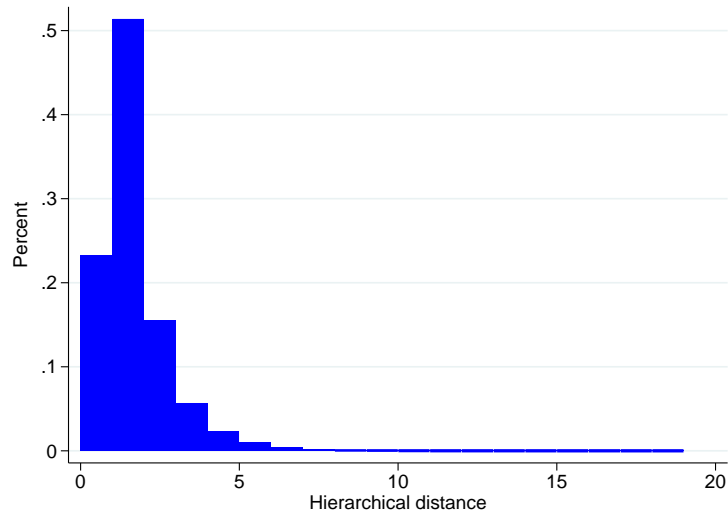


Figure 1: Frequency of hierarchical distances in firm-year observations

In contrast to the other two business group controls discussed above, hierarchical

⁸Note that $[\mathbf{I} - \mathbf{R}]^{-1}$ is commonly known as the Leontief inverse matrix.

distance varies over firms within the same business group. From Table 1, we see that both the mean and the standard deviation of our hierarchical distance variable are fairly small, indicating a concentration of the firm-year observations at the bottom of its domain. Figure 1 shows the hierarchical distances observed in our dataset. Since H_j is a continuous variable, we report its frequencies for symmetric intervals of unit length. The first bar of the histogram captures the 23.2 percent ultimate owners among our firm-year observations. The residual bars refer to the 76.8 percent subsidiaries, for which we find a strongly right-skewed density. Thus, a significant fraction of firm-year observations show a low hierarchical distance to their ultimate owner.

Establishment and worker variables from IEB

Table 2 reports key summary statistics for the establishments and workers linked from the Integrated Employment Biographies (IEB) to the Orbis data. We count 21,609,088 worker-year observations that can be aggregated to 430,699 establishment-year observations over the observation period 2013–2017. The variation in log establishment size is fairly high and a major part of establishments come from three broad sector categories, namely manufacturing, retail & repair, and finance & insurance. Moreover, we observe considerable variation in (imputed) log daily wages, sizable age differences of workers, and an underrepresentation of females. Classifying workers with no vocational training and no high-school degree as low-skilled, workers with a high-school degree and/or vocational training as medium-skilled, and workers with a degree from a university or a university of applied sciences as high-skilled, we find strong differences in the coverage of skill groups, with medium-skilled workers accounting for more than 70 percent of the worker-year observations.

3.2 Hierarchical distance and wages

Before turning to the econometric analysis, we use the linked IEB-Orbis dataset to provide descriptive evidence on how hierarchical distance to the ultimate owner affects workers' wages. To cancel out the impact of other covariates that have shown to be important wage determinants by previous empirical research, we first run a Mincer (1958)-type regression, in which we explain the log (daily) wage by worker observables on age (as a proxy for experience), age squared, and dummies for three skill groups, German nationality, female gender, and 16 federal states. We additionally control for time dummies and the six broad sector categories listed in

Table 2: Establishment and worker characteristics

	<i>Mean</i>	<i>Std. Dev.</i>
<i>(a) Establishment characteristics</i>		
Log employment	2.997	1.497
Agriculture	0.008	0.087
Manufacturing	0.161	0.368
Mining, utilities & construction	0.080	0.271
Retail & repair	0.444	0.497
Finance & insurance	0.211	0.408
Private & public services	0.096	0.295
<i>(b) Worker characteristics</i>		
Log wage	4.815	0.496
Age	42.6	11.2
Female	0.265	0.442
Low-skilled	0.048	0.046
Medium-skilled	0.716	0.451
High-skilled	0.236	0.425

Notes: Establishment and worker descriptives are constructed for the years 2013-2017, using Integrated Employment Biographies (IEB) from the Institute for Employment Research in Nuremberg. Establishment characteristics are computed for 430,699 establishment-year observations. Worker characteristics are computed for 21,609,088 worker-year observations. Low-skilled workers have no vocational training and no high-school degree. Workers with a high-school degree and/or vocational training are medium-skilled, whereas workers holding a degree from a university or a university of applied sciences are high-skilled.

Table 2. To illustrate the correlation between the residual wage of workers and their hierarchical distance to the ultimate owner, we assign establishments the hierarchical distance we have computed for the firm they are merged with, cluster establishments into deciles of hierarchical distances, and compute averages of hierarchical distances and residual wages for these deciles. We then plot each pair of averages as individual data point in Figure 2.⁹

Figure 2 shows a positive relationship between the hierarchical distance to the ultimate owner of a business group and workers' wages. According to our theoretical model outlined in Section 2 this indicates that monitoring efficiency decreases with larger hierarchical distance to the ultimate owner, making it necessary to incentivize workers employed in hierarchically distant establishments by paying higher

⁹Since a high frequency of firms (and thus establishments) show a hierarchical distance of zero or one, the number of displayed data points is less than 10.

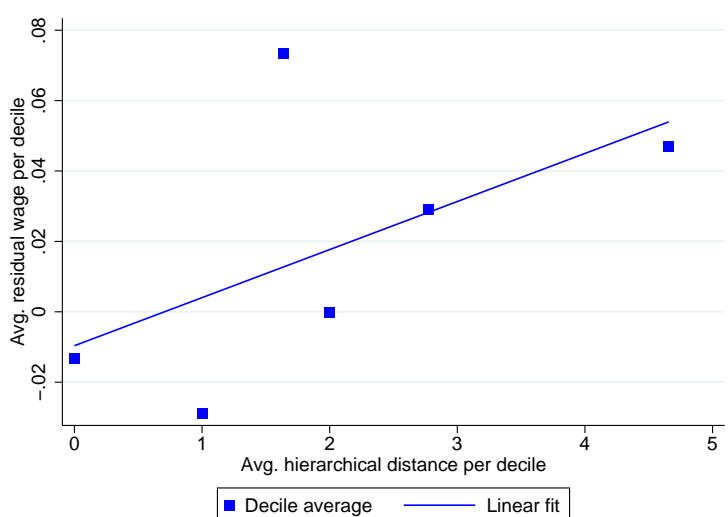


Figure 2: Hierarchical distance and wages in business groups

wages. However, the evidence reported in Figure 2 is far from being conclusive and does not allow for causal inference on how changes in hierarchical distance affect workers' wages. In the empirical analysis of Section 4, we analyze the link between hierarchical distance and wage payments in business groups in further detail.

4 Estimation and empirical results

To study the role of business groups for individual wages in a systematic way, we first run OLS and fixed-effects regressions, in which we control for observable and unobservable worker and establishment characteristics. To avoid selection bias, we use in a second step propensity-score matching and select a control group that is (ex ante) comparable to our treatment group. We then determine the effect of changes in hierarchical distance to the ultimate owner on wages by a difference-in-difference approach.

4.1 Baseline estimations

In the subsequent analysis, we estimate a model of the following form:

$$w_{ijkt} = \alpha + \mathbf{X}_{it} \cdot \beta + \mathbf{C}_{jt} \cdot \gamma + \mathbf{N}_{jkt} \cdot \nu + \mu_t + \epsilon_{ijkt}, \quad (9)$$

where w_{ijkt} is the log daily wage of worker i in establishment j , business group k , and year t and α is a constant. \mathbf{X}_{it} is a (row) vector of the (time-varying) worker covariates age, age squared, and dummies for three skill groups, German nationality, and female gender, with β as the corresponding (column) vector of coefficients. \mathbf{C}_{jt} is a vector of the (time-varying) establishment covariates log employment, log employment squared, and dummies for 15 German federal states and five broad sector categories, with γ as the corresponding vector of coefficients. Moreover, \mathbf{N}_{jkt} is a vector of business group determinants, which include as the main variable of interest the hierarchical distance index and vary across establishments, business groups, and time, with ν as the respective vector of coefficients. Finally, μ_t is a vector of time dummies and ϵ_{ijkt} is the error term.

Whereas the baseline specification in Eq. (9) determines a flexible model that allows for non-linear effects, the estimated coefficients are prone to omitted variable bias if our set of controls does not cover all important worker, establishment, and business group determinants of wages. We capture unobserved, time-invariant determinants by adding worker-establishment-(business-)group fixed-effects. This gives a modified regression model of the following form:

$$w_{ijkt} = \alpha + \mathbf{X}_{it} \cdot \beta + \mathbf{C}_{jt} \cdot \gamma + \mathbf{N}_{jkt} \cdot \nu + \mu_t + \phi_{ijk} + \epsilon_{ijkt}, \quad (9')$$

where ϕ_{ijk} denotes worker-establishment-group fixed-effects. By including these fixed-effects, we time-demean each worker-establishment-group observation and identify the effects of changes in the business group covariates through their variation over time. A change in the hierarchical distance variable can then only exert an effect on wages if a worker-establishment observation changes its hierarchical position in a given business group (by adding or dropping hierarchical layers). However, the effects arising from time-invariant worker, establishment, and business group determinants as well as the effects of workers switching the establishment or of establishments switching the business group are eliminated. This allows us to isolate the effect of changes in hierarchical distance from other factors influencing workers' wages, such as firm size or foreign ownership wage premia. As a result, the regression model in Eq. (9') gives consistent estimates of ν , but it may underestimate the overall importance of business group variables for wages in our dataset.

Table 3 shows our estimation results. In all regressions, we control for the full set of worker and establishment covariates reported in Table 2 and additionally include time and federal state dummies. In Models (1), (3), and (5) we estimate

Eq. (9) using OLS, whereas the remaining models refer to fixed-effects regressions based on Eq. (9'). Model (1) captures the most parsimonious specification and only includes the hierarchical distance of establishments to the ultimate owner as a business group control. The estimated effect is positive and significant at the one percent level. Increasing the hierarchical distance by one standard deviation ($\cong 1.20$) increases wages by 1.71 log points. Abstracting from circular ownership linkages, one can interpret the size of this effect in a more intuitive way as follows. Moving down one layer in business group hierarchy would increase worker's wages by almost two percent. Model (2) shows that the size of this effect decreases when controlling for worker-establishment-group fixed-effects.

Table 3: Business groups, ownership hierarchy, and wages

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
log daily wage	OLS	FE	OLS	FE	OLS	FE
Hierarchical distance	0.0145*** (0.0014)	0.0013* (0.0007)	0.0130*** (0.0015)	0.0028*** (0.0008)	0.0025 (0.0026)	0.0040*** (0.0005)
Group size			0.0103*** (0.0007)	0.0013*** (0.0003)		
Hierarchical distance × Group size			-0.0012*** (0.0002)	-0.0002*** (0.0001)		
Group complexity					0.0269*** (0.0013)	0.0015* (0.0008)
Hierarchical distance × Group complexity					-0.0030*** (0.0003)	-0.0005*** (0.0002)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Worker Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Establishment Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Worker-establ.-group FE	No	Yes	No	Yes	No	Yes
R-sq. (within)	0.4440	0.0740	0.4504	0.0740	0.4500	0.0740
Observations	21,609,088	21,609,088	21,609,088	21,609,088	21,609,088	21,609,088

Notes: Worker covariates include age, age squared, dummies for two skill groups, German nationality, and gender. Establishment covariates include log employment, log employment squared, dummies for 15 German federal states, five broad sector categories, and four time dummies. In all models, we estimate a constant as well as time dummies. Hierarchical distance and the group index of complexity are constructed as outlined in Section 3. Standard errors in parentheses are clustered at the establishment-level. ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

In Models (3) to (6) we add further business group covariates. In Models (3) and (4) these are group size as well as its interaction with hierarchical distance. Adding these controls has rather small effects on our hierarchical distance estimate. Moreover, the positive direct effect of group size is well in line with evidence on

size-wage premia at the firm level (cf. Colonelli et al., 2018). The negative sign of the interaction term indicates that hierarchical distance is less important for wages in larger business groups. In Models (5) and (6) we consider the role of group complexity and its interaction with hierarchical distance. In the OLS regression we find that the impact of hierarchical distance, while staying positive, becomes considerably smaller than in the parsimonious specification of Model (1) and loses its statistical significance. In contrast, the direct effect of higher group complexity on wages is positive, sizable, and significant. This result changes drastically in the fixed-effects regression. Controlling for unobserved time-invariant heterogeneity of workers, establishments, and business groups, we find a positive and significant impact of larger hierarchical distance to the ultimate owner on workers' wages, while the impact of group complexity falls considerably. The negative and significant interaction term indicates that hierarchical distance plays a less important role for wages in more complex business groups.

Table 4 reports results, distinguishing between low-, medium-, and high-skilled workers. Thereby, we rely on our preferred specifications from Models (5) and (6), where we have added group complexity to separate the impact of differences in the horizontal and vertical dimension of business groups from the hierarchical distance effect. By and large the insights from the pooled sample in Table 3 extend to the sub-groups of low- and medium-skilled workers. There is no evidence for a positive effect of hierarchical distance on wages when relying on OLS. However, once controlling for unobserved heterogeneity of workers, establishment, and business groups by fixed-effects, larger hierarchical distance to the ultimate owner of the business group exerts a positive effect on wages paid by an establishment to its low- and medium-skilled workers. Moreover, for these two sub-groups of workers the direct effect of group complexity on wages becomes considerably weaker, whereas the interaction between hierarchical distance and group complexity remains negative and significant. However, for the sub-group of high-skilled workers, for which the effects of the business group variables are strong with OLS, but disappear when accounting for worker-establishment-group fixed-effects.

Summing up, the results from Tables 3 and 4 show that the omitted variable bias in estimating the link between hierarchical distance to the ultimate owner of a business group and workers' wages with simple OLS can be severe, so that controlling for unobserved heterogeneity appears highly important. Moreover, the results from fixed-effects regressions are broadly in line with the descriptive evidence and support the conclusion that larger hierarchical distance reduces monitoring efficiency, so

Table 4: Business groups, ownership hierarchy, and skill-specific wages

Dependent variable:	Low skilled		Medium skilled		High skilled	
	(1)	(2)	(3)	(4)	(5)	(6)
log daily wage	OLS	FE	OLS	FE	OLS	FE
Hierarchical distance	-0.0072* (0.0040)	0.0063*** (0.0022)	0.0011 (0.0027)	0.0048*** (0.0015)	0.0104*** (0.0031)	-0.0002 (0.0013)
Group complexity	0.0264*** (0.0021)	0.0009 (0.0010)	0.0266*** (0.0014)	0.0020* (0.0010)	0.0271*** (0.0014)	0.0002 (0.0006)
Hierarchical distance × Group complexity	-0.0025*** (0.0006)	-0.0004* (0.0003)	-0.0028*** (0.0004)	-0.0006*** (0.0002)	-0.0037*** (0.0003)	-0.0001 (0.0001)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Worker Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Establishment Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Worker-establ.-group FE	No	Yes	No	Yes	No	Yes
R-sq. (within)	0.4266	0.1188	0.3290	0.0901	0.3286	0.0529
Observations	1,036,435	1,036,435	15,471,201	15,471,201	5,101,452	5,101,452

Notes: Worker covariates include age, age squared, dummies for two skill groups, German nationality, and gender. Establishment covariates include log employment, log employment squared, dummies for 15 German federal states, five broad sector categories, and four time dummies. In all models, we estimate a constant as well as time dummies. Hierarchical distance and the group index of complexity are constructed as outlined in Section 3. Standard errors in parentheses are clustered at the establishment-level. ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

that higher wages are needed to incentivize workers from establishments at low hierarchical layers of the business group.

Whereas fixed-effects regressions are an effective remedy for time-invariant omitted variable bias, our estimates may still be prone to a selection bias that exists, for instance, if the position of subsidiaries in business group hierarchy is not random. To rule out a selection bias, we exploit in the next subsection a two-stage regression procedure, combining propensity-score matching with a difference-in-difference approach.

4.2 Selection into business groups

In line with our analysis in Section 4.1, we specify the treatment as an increase in the hierarchical distance (HD) between the ultimate owner and a worker-establishment pair. To isolate the hierarchical distance effect from other wage determinants associated with employer effects, we focus on workers who stay within the same establishment and business group around the treatment event. Therefore, we define the treatment as an increase in hierarchical distance between a worker-establishment

pair and its ultimate owner within a given business group. Accordingly, we classify worker-establishment pairs as untreated if they show either a decline or no change of hierarchical distance to their ultimate owner within a given business group.

Following the matching literature, we collapse the observation period 2013 to 2017 into two-year windows around the treatment period and eliminate all observations that are not classified as treated or untreated according to the definition introduced above. Moreover, being interested in the *average* wage effects for individual workers, we specify the treatment at the worker level and capture this treatment by a binary indicator

$$D_{ijk} = \begin{cases} 1 & HD_{ijk,t=0} < HD_{ijk,t=1} \\ 0 & HD_{ijk,t=0} \geq HD_{ijk,t=1} \end{cases}, \quad (10)$$

which takes a value of one, if the hierarchical distance to the ultimate owner of worker i from establishment j and business group k increases between periods $t = 0$ and $t = 1$. In contrast, the indicator takes a value of zero if the hierarchical distance to the ultimate owner of worker i from establishment j and business group k does not increase. Making use of our narrow definitions, we end up with a sample of 20,882,842 treated and untreated worker-year observations.¹⁰

To select for each treated observation a suitable control from the pool of untreated worker-establishment pairs, we rely on nearest-neighbour propensity-score matching (Rosenbaum and Rubin, 1983). For this purpose, we determine the probability in $t = 0$ that an observation is subject to treatment between periods 0 and 1 and estimate the following probit model:

$$P(D_{ijk} = 1) = \Phi(\nu \cdot \mathbf{N}_{jk,0} + \gamma \cdot \mathbf{C}_{j,0} + \beta \cdot \mathbf{X}_{i,0}), \quad (11)$$

where $\mathbf{N}_{jk,0}$, $\mathbf{C}_{j,0}$, $\mathbf{X}_{i,0}$ are vectors of business group, establishment, and worker covariates in period $t = 0$, with ν , γ and β being the corresponding vectors of coefficients. Business group covariates are hierarchical distance and group complexity. Establishment covariates include the log of employment to control for establishment size, sector dummies indicating the establishments' industry affiliation, and federal

¹⁰Since we observe each worker for at least two consecutive years, our sample contains 4,811,509 unique worker observations – with 597,638 of them showing an increase in the hierarchical distance between the first and the second observation year, thereby forming our treatment group. For the residual 4,213,871 workers we observe either a decline or no change in the hierarchical distance to their ultimate owner.

state dummies to control for establishment location. Finally, worker covariates cover dummies for females and three skill levels, workers' age, and their log daily wages. To exclude time effects, we estimate the propensity score (11) within treatment cohorts, i.e. we match observations from the same year.¹¹ Since six observations are off support, we eliminate them from our treatment group after the probit estimation.

Using the estimates from our probit model, we then assign to each worker-establishment-group triple from the treatment group the worker-establishment-group triple from the pool of untreated observations with the smallest absolute difference in the propensity-score. This forms our control group, which contains fewer unique observations than the treatment group, because we match with replacement (see Caliendo and Kopeinig, 2008). Moreover, since we match on individuals, workers from a single establishment of the treatment group can be assigned to workers from different establishments belonging to different ownership networks in the control group.

To evaluate the success of our matching procedure, we compare averages of all covariates used in the probit estimation before and after matching and report the results in the Appendix. There, we discuss two diagnostics that are commonly used to assess the matching quality. The first one is the standardized percentage bias introduced by Rosenbaum and Rubin (1985). Matching reduces the mean bias considerably from 12.1 percent to 1.9 percent. We also report the normalized difference between covariates, as put forward by Imbens and Wooldridge (2009). Imbens and Wooldridge (2009) suggest an upper limit of one quarter to consider a variable as balanced. This critical threshold is not surpassed by any of our covariates after matching. The two diagnostics therefore indicate that we were successful in matching observations from the treatment group to similar untreated observations.

With the matched sample at hand, we can quantify the causal effect of larger hierarchical distance to the ultimate owner on wages using a difference-in-difference approach. In doing so, we contrast wages before and after treatment and compare the change in wages between workers from the treatment and the control group by estimating the following equation:

$$w_{ijkt} = \alpha_i + \mu + \eta \cdot D_{ijk} + \epsilon_{ijkt}, \quad (12)$$

¹¹As a robustness check, we account for the change in log employment prior to treatment as an additional control in the probit model. We include this variable to take employment dynamics prior to the treatment into account and report the results along with those from three further robustness checks in the Appendix.

where w_{ijkt} is the log daily wage of worker i in establishment j , business group k , and year t , α_i is a worker fixed-effect to control for any remaining, time-invariant unobserved heterogeneity of workers, and μ is a time dummy that takes a value equal to one in the post-treatment period $t = 1$. D_{ijk} is the treatment indicator equal to one for each stayer i , whose establishment j has been subject to treatment between $t = 0$ and $t = 1$, and zero otherwise. Coefficient η captures the wage effect for workers, whose establishment increases its hierarchical distance to the ultimate owner within a given business group. Finally, ϵ_{ijkt} is the error term.

Table 5: Wage effect of larger distance in business group hierarchy

Dependent variable:	<i>All workers</i>	<i>Low-skilled</i>	<i>Med.-skilled</i>	<i>High-skilled</i>
Log daily wage				
Higher <i>HD</i> in $t = 1$	0.0130*** (0.0013)	0.0140*** (0.0025)	0.0105*** (0.0014)	0.0197*** (0.0017)
Observations	2,390,528	120,276	1,680,908	589,344

Notes: The treatment is defined as an increase in the hierarchical distance within a given business group. The estimation includes a time dummy and worker fixed-effects. Standard errors in parentheses are clustered at the establishment-level. ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

Table 5 summarizes the estimation results for the pooled sample of all workers as well as for the three skill groups. In line with the fixed-effects regressions in Section 4.1, we find for the pooled sample that larger hierarchical distance to the ultimate owner of the business group increases workers' wages by 1.3 percent. This picture is unchanged when zooming in on the three skill groups, with a positive and significant coefficient of the treatment indicator found for all sub-groups.

Overall, the results from Table 5 support to the prediction of our theoretical model that monitoring efficiency decreases in hierarchical distance, inducing the ultimate owner to increase incentive pay for workers employed by establishments at comparably low layers in business group hierarchy.

In a final step, we investigate whether our results are robust to changes in our treatment definition using the full sample of all workers. In Table 6, Model(1), we set a threshold for the hierarchical distance of 0.25 and drop observations with changes in hierarchical distance smaller than this threshold from the treatment group. With this refinement, we eliminate changes in ownership shares that are too small to induce relevant adjustments in hierarchical control. Table 6 shows that introducing a lower threshold for the hierarchical distance variable reduces the treatment effect,

but does not change our results in a substantive way.¹² In Model (2), we drop all observations showing an increase in hierarchical distance larger than two. This makes the treatment group more homogeneous and ensures that our results are not driven by a small number of outliers. Introducing the upper bound has no effect on the estimation result.

Table 6: Hierarchical distance and wages: alternative specifications

Dependent variable:	<i>All workers</i>			
Log daily wage	Model (1)	Model (2)	Model (3)	Model (4)
Higher <i>HD</i> in $t = 1$	0.0090*** (0.0016)	0.0130*** (0.0013)	0.0065*** (0.0022)	0.0057*** (0.0007)
Observations	1,868,964	2,343,304	2,998,600	4,876,312

Notes: In Model (1), we confine the treatment to increases in hierarchical distance by at least 0.25. In Model (2), we confine the treatment to increases in hierarchical distance by at most two. In Model (3), we consider worker-establishment pairs that change their business group around the treatment period. In Model (4), we broaden the definition of treatment and control group, including worker-establishment pairs that stay in their business group as well as worker-establishment pairs that change their business group. All estimations include a time dummy and worker fixed-effects. Standard errors in parentheses are clustered at the establishment-level. ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

In two further exercises we do no longer restrict the analysis to worker-establishment observations that stay in the same business group around the treatment event. In Model (3), we keep only worker-establishment pairs that are subject to ownership change around the treatment period. This lowers the treatment effect by 50 percent. In Model (4), we do not condition on takeover events and include all worker-establishment observations in our analysis [EJ hier ist mir nicht klar, was gemacht wurde. In dem Fall müssten wir ja über 20 Mio Beobachtungen haben. Sind das wieder nur zwei Perioden?]. Similar to Model (3), changing the definition of treated and untreated observations lowers the treatment effect considerably. Of course, choosing the alternative definitions, we cannot rule out that the treatment effect captures at least partially the impact of ownership change, which has been put forward to exert sizable wage effects in the context of multinational enterprises. However, ownership change is not confined to foreign takeover in our analysis, so that the sizable drop of the treatment effect does not contradict existence of a foreign

¹²Increasing the threshold to 0.5 or 0.75 would further reduce sample size but not change our results qualitatively.

ownership wage premium.¹³

5 Conclusion

In this paper, we merge firm-level data on ownership linkages in business groups with administrative worker and establishment data for Germany to create a new dataset, which allows to analyze how the position in business group hierarchy affects workers' wages. Since the ownership linkages are not one-directional, we propose a measure of hierarchical distance that acknowledges the complex structure of business groups in our data. In our baseline regressions we find clear evidence for a positive impact of larger hierarchical distance on wages for the pooled sample as well as for the sub-groups of low- and medium skilled workers. For high-skilled workers the link between the position in business group hierarchy and workers' wages is less clear.

To shield our estimates from selection bias, we consider in a further step a two-stage estimation approach, in which we first select a treatment and control group based on propensity-score matching and then estimate the treatment effect of an increase in the hierarchical distance to the ultimate owner on workers' wages using a difference-in-difference estimator. The results from this more elaborate empirical approach are similar to those from OLS and fixed-effects regressions. Larger hierarchical distance exerts a positive effect on workers' wages, with the effect being remarkably robust to changes in the composition of the treatment group.

Overall, our results speak for a sizable impact of larger hierarchical distance to the ultimate owner of a business group on workers' wages. In a parsimonious specification, we show that increasing hierarchical distance by one standard deviation or approximately one layer increases wages by almost two log points. The empirical evidence reported in this paper is consistent with a monitoring-based theory of business group hierarchies, in which larger hierarchical distance to the ultimate owner is associated with lower monitoring efficiency and higher wages for incentivizing workers.

¹³In two extensions to Models (3) and (4), we have added a dummy for foreign takeover and its interaction term with the treatment indicator. In these extensions, which are available upon request, we find evidence for both a positive wage effect of larger hierarchical distance and a positive wage effect of foreign takeover.

A Appendix

A.1 Balancing test for the matching procedure

Table A.1: Balancing test for the matching procedure with replacement

<i>Variable</i>	<i>Sample</i>	<i>Mean</i>		<i>Stand.</i>	<i>Bias</i>	<i>Normal.</i>
		<i>Treated</i>	<i>Control</i>	<i>bias %</i>	<i>reduction</i>	<i>diff.</i>
<i>(a) Group characteristics</i>						
Hierarchical distance	Unmatched	2.313	1.112	102.6		
Hierarchical distance	Matched	2.313	2.302	1.0	99.0	0.006
Group complexity	Unmatched	4.173	2.151	74.3		
Group complexity	Matched	4.173	4.023	5.5	92.6	0.035
<i>(b) Establishment characteristics</i>						
Log employment	Unmatched	6.010	5.993	0.9		
Log employment	Matched	6.010	5.928	4.5	-397.3	0.033
Agriculture	Unmatched	0.000	0.003	-7.0		
Agriculture	Matched	0.000	0.000	0.2	97.4	0.004
Manufacturing	Unmatched	0.538	0.454	17.0		
Manufacturing	Matched	0.538	0.535	0.6	96.2	0.005
Mining, util. & constr.	Unmatched	0.059	0.080	-8.5		
Mining, util. & constr.	Matched	0.059	0.053	2.3	73.3	0.018
Retail & repair	Unmatched	0.211	0.240	-7.0		
Retail & repair	Matched	0.211	0.217	-1.5	78.9	-0.011
Finance & insurance	Unmatched	0.164	0.122	12.1		
Finance & insurance	Matched	0.164	0.182	-5.1	58.0	-0.033
Priv. & publ. services	Unmatched	0.028	0.101	-30.1		
Priv. & publ. services	Matched	0.028	0.013	6.1	79.8	0.074
Schleswig-Holstein	Unmatched	0.023	0.021	1.3		
Schleswig-Holstein	Matched	0.023	0.023	-0.2	81.5	-0.002
Hamburg	Unmatched	0.053	0.030	11.3		
Hamburg	Matched	0.053	0.048	2.3	80.0	0.015
Lower Saxony	Unmatched	0.080	0.086	-2.2		
Lower Saxony	Matched	0.080	0.074	2.2	-2.0	0.016
Bremen	Unmatched	0.015	0.012	2.0		
Bremen	Matched	0.015	0.012	2.4	-19.0	0.017
North Rhine-Westphalia	Unmatched	0.200	0.201	-0.2		
North Rhine-Westphalia	Matched	0.200	0.205	-1.1	-362.2	-0.008
Hesse	Unmatched	0.123	0.082	13.7		
Hesse	Matched	0.123	0.130	-2.2	83.9	-0.014
Rhineland-Palatinate	Unmatched	0.035	0.044	-4.6		
Rhineland-Palatinate	Matched	0.035	0.034	0.6	86.7	0.005
Baden-Württemberg	Unmatched	0.138	0.166	-7.9		
Baden-Württemberg	Matched	0.138	0.137	0.4	94.7	0.003
Bavaria	Unmatched	0.170	0.176	-1.5		
Bavaria	Matched	0.170	0.177	-1.9	-29.1	-0.013
Saarland	Unmatched	0.006	0.015	-8.2		
Saarland	Matched	0.006	0.008	-2.0	75.3	-0.017
Berlin	Unmatched	0.031	0.036	-2.8		
Berlin	Matched	0.031	0.033	-1.1	62.1	-0.008
Brandenburg	Unmatched	0.028	0.023	3.6		

Table A.1 – *continued from previous page*

<i>Variable</i>	<i>Sample</i>	<i>Mean</i>		<i>Stand.</i>	<i>Bias</i>	<i>Normal.</i>
		<i>Treated</i>	<i>Control</i>	<i>bias %</i>	<i>reduction</i>	<i>diff.</i>
Brandenburg	Matched	0.028	0.029	-0.2	93.1	-0.002
Mecklenburg West-Pomerania	Unmatched	0.012	0.013	-0.7		
Mecklenburg West-Pomerania	Matched	0.012	0.010	1.8	-155.5	0.013
Saxony	Unmatched	0.044	0.049	-2.6		
Saxony	Matched	0.044	0.041	1.4	46.4	0.010
Saxony-Anhalt	Unmatched	0.026	0.022	2.4		
Saxony-Anhalt	Matched	0.026	0.025	0.5	77.6	0.004
Thuringia	Unmatched	0.016	0.024	-5.6		
Thuringia	Matched	0.016	0.015	0.8	86.2	0.006
<i>(c) Worker characteristics</i>						
Female	Unmatched	0.229	0.252	-5.3		
Female	Matched	0.229	0.227	0.7	87.4	0.005
Age	Unmatched	42.9	42.8	0.6		
Age	Matched	42.9	42.8	0.5	7.3	0.004
Low skilled	Unmatched	0.051	0.045	2.5		
Low skilled	Matched	0.051	0.052	-0.5	81.5	-0.003
Medium skilled	Unmatched	0.703	0.739	-8.0		
Medium skilled	Matched	0.703	0.688	3.3	58.5	0.023
High skilled	Unmatched	0.246	0.216	7.2		
High skilled	Matched	0.246	0.260	-3.3	54.5	-0.023
Log wage	Unmatched	4.915	4.820	20.2		
Log wage	Matched	4.915	4.932	-3.6	82.4	-0.025
<i>Sample</i>				<i>Mean bias</i>	<i>Median bias</i>	
Unmatched				12.1	5.6	
Matched				1.9	1.5	

Notes: All variables are measured in $t = 0$ and averaged at the worker-level in the treated and control group respectively.

A.2 Further robustness checks

To make sure that the positive effect of larger hierarchical distance on wages reported in Section 4.2 is robust to different specifications of the propensity-score matching, we report in Table A.2 the results for the pooled sample of all workers, relying on four alternatives to our main matching procedure. In Model (1), we match without replacement and find that this does not change the treatment effect. In Model (2), we add the difference in log establishment employment between period $t = -1$ and $t = 0$ as a further covariate in the probit model. This allows us to control for differences in the employment dynamics prior to the treatment. Adding this covariate somewhat reduces sample size and slightly lowers the treatment effect, while leaving unchanged the main insight from our baseline specification in Table 5. In Model (3), we replace in the probit model the continuous log employment variable by dummies for five

establishment size categories. Thereby, we distinguish establishments with less than ten, between ten and 49, between 50 and 249, between 250 and 999, and with more than 1000 employees. Additionally, we control for employment dynamics prior to treatment by introducing two dummies equal to one if the establishment has either increased or decreased its workforce by at least three percent between $t = -1$ and $t = 0$ (with the omitted category referring to establishments with an absolute change in workforce size by less than three percent). This modification has a comparably small effect on the estimated treatment effect.

Table A.2: The effect of an increase in hierarchical distance on wages

Dependent variable:	<i>All workers</i>			
Log daily wage	Model (1)	Model (2)	Model (3)	Model (4)
Higher <i>HD</i> in $t = 1$	0.133*** (0.0012)	0.0121*** (0.0014)	0.0122** (0.0014)	0.0053*** (0.0020)
Observations	2,390,528	2,373,652	2,390,492	2,517,208

Notes: In Model (1), we match without replacement. In Model (2) we take into account employment dynamics prior to treatment by additionally matching on the difference in log (establishment) employment between $t = -1$ and $t = 0$. In Model (3), we match on five establishment size categories and two dummy variables indicating an absolute change in log (establishment) employment between $t = -1$ and $t = 0$ of at least three percent. In Model (4), we define the treatment at the establishment level and match accordingly. All estimations include a time dummy and worker fixed-effects. Standard errors in parentheses are clustered at the establishment-level. ***, ** and * denote significance at the 1, 5 and 10 percent levels, respectively.

In the final robustness check of Model (4), we define the treatment at the establishment and not the worker level, thereby imposing the constraint that all workers from a given establishment in the treatment group are matched with workers from a single establishment of the control group. As expected, imposing the additional constraint lowers matching quality. Moreover, it reduces the estimated treatment effect by more than 50 percent, while leaving the main insight from our empirical analysis that larger hierarchical distance increases wages intact.

References

- ALTOMONTE, C., G. OTTAVIANO, AND A. RUNGI (2018): “Business Groups as Knowledge-Based Hierarchies of Firms,” *Working Paper*.
- ALTOMONTE, C. AND A. RUNGI (2013): “Business groups as hierarchies of firms: determinants of vertical integration and performance,” *ECB Working Paper 1554*.
- (2015): “Firm Performance and the Optimal Organization of Business Groups,” *Working Paper*.
- ANTONI, M., K. KOLLER, M.-C. LAIBLE, AND F. ZIMMERMANN (2018): “Orbis-ADIAB: From record linkage key to research dataset,” *FDZ Methodenreport 04/2018*.
- ANTRÀS, P. AND D. CHOR (2013): “Organizing the Global Value Chain,” *Econometrica*, 81, 2127–2204.
- ANTRÀS, P., D. CHOR, T. FALLY, AND R. HILLBERRY (2012): “Measuring the Upstreamness of Production and Trade Flows,” *American Economic Review: Papers and Proceedings*, 102, 412–16.
- ATALAY, E., A. HORTAÇSU, AND C. SYVERSON (2014): “Vertical Integration and Input Flows,” *American Economic Review*, 104, 1120–1148.
- BAKER, G., R. GIBBONS, AND K. J. MURPHY (2002): “Relational Contracts and the Theory of the Firm,” *The Quarterly Journal of Economics*, 117, 39–84.
- BALSVIK, R. AND S. A. HALLER (2010): “Picking Lemons or Picking Cherries? Domestic and Foreign Acquisitions in Norwegian Manufacturing,” *The Scandinavian Journal of Economics*, 112, 361–387.
- BASTOS, P., N. MONTEIRO, AND O. R. STRAUME (2018): “Foreign Acquisition and Internal Organization,” *Journal of International Economics*, 114, 143–163.
- BLOOM, N., F. GUVENEN, B. S. SMITH, J. SONG, AND T. VON WACHTER (2018): “The Disappearing Large-Firm Wage Premium,” *AEA Papers and Proceedings*, 108, 317–22.
- BLOOM, N., R. SADUN, AND J. VAN REENEN (2010): “Recent Advances in the Empirics of Organizational Economics,” *Annual Review of Economics*, 2, 105–137.
- BROWN, C. AND J. MEDOFF (1989): “The Employer Size-Wage Effect,” *Journal of Political Economy*, 97, 1027–1059.
- CALIENDO, L., F. MONTE, AND E. ROSSI-HANSBERG (2015): “The Anatomy of French Production Hierarchies,” *Journal of Political Economy*, 123, 809–852.
- CALIENDO, L. AND E. ROSSI-HANSBERG (2012): “The Impact of Trade on Organization and Productivity,” *The Quarterly Journal of Economics*, 127, 1393–1467.

- CALIENDO, M. AND S. KOPEINIG (2008): “Some Practical Guidance For The Implementation Of Propensity Score Matching,” *Journal of Economic Surveys*, 22, 31–72.
- CALVO, G. A. AND S. WELLISZ (1978): “Supervision, Loss of Control, and the Optimum Size of the Firm,” *Journal of Political Economy*, 86, 943–952.
- (1979): “Hierarchy, Ability, and Income Distribution,” *Journal of Political Economy*, 87, 991–1010.
- CARD, D., J. HEINING, AND P. KLINE (2013): “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *The Quarterly Journal of Economics*, 128, 967–1015.
- CAROLI, E. AND J. VAN REENEN (2001): “Skill-biased Organizational Change? Evidence from a Panel of British and French Establishments,” *The Quarterly Journal of Economics*, 116, 1449–1492.
- CHEN, C. (2017): “Management Quality and Firm Hierarchy in Industry Equilibrium,” *American Economic Journal: Microeconomics*, 9, 203–244.
- CHEN, C. AND W. SUEN (2019): “The Comparative Statics of Optimal Hierarchies,” *American Economic Journal: Microeconomics*, 11, 1–25.
- COLONELLI, E., J. TÅG, M. WEBB, AND S. WOLTER (2018): “A Cross-Country Comparison of Dynamics in the Large Firm Wage Premium,” *American Economic Review: Papers and Proceedings*, 108, 323–327.
- COSTINOT, A., J. VOGEL, AND S. WANG (2012): “An Elementary Theory of Global Supply Chains,” *The Review of Economic Studies*, 80, 109–144.
- DAUTH, W. AND J. EPPELSHEIMER (2020): “Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide,” *Journal for Labor Market Research*, 54.
- DUSTMANN, C., B. FITZENBERGER, U. SCHÖNBERG, AND A. SPITZ-OENER (2014): “From Sick Man of Europe to Economic Superstar: Germany’s Resurgent Economy,” *Journal of Economic Perspectives*, 28, 167–188.
- EGGER, H. AND E. JAHN (2020): “Ownership and the multinational wage premium,” *Applied Economics Letters*, 27.
- EGGER, H., E. JAHN, AND S. KORNITZKY (2020): “Reassessing the foreign ownership wage premium in Germany,” *The World Economy*, 43, 302–325.
- FRIEDRICH, B. U. (forthcoming): “Trade Shocks, Firm Hierarchies, and Wage Inequality,” *Review of Economics and Statistics*.

- GARICANO, L. (2000): “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 108, 874–904.
- GARICANO, L. AND E. ROSSI-HANSBERG (2006): “Organization and Inequality in a Knowledge Economy,” *The Quarterly Journal of Economics*, 121, 1383–1435.
- GIRMA, S. AND H. GÖRG (2007): “Evaluating The Foreign Ownership Wage Premium Using a Difference-in-Differences Matching Approach,” *Journal of International Economics*, 72, 97–112.
- GIRMA, S., D. GREENAWAY, AND K. WAKELIN (2001): “Who Benefits from Foreign Direct Investment in the UK?” *Scottish Journal of Political Economy*, 48, 119–133.
- GUMPERT, A. (2018): “The organization of knowledge in multinational firms,” *Journal of the European Economic Association*, 16, 1929–1976.
- HANSON, GORDON, H., R. J. J. TATALON, AND M. J. SLAUGHTER (2005): “Vertical Production Networks in Multinational Firms,” *Review of Economics and Statistics*, 87, 664–678.
- HIJZEN, A., P. S. MARTINS, T. SCHANK, AND R. UPWARD (2013): “Foreign owned firms around the world: A comparative analysis of wages and employment at the micro-level,” *European Economic Review*, 60, 170–188.
- HILL, N. S., M.-G. SEO, J. H. KANG, AND M. S. TAYLOR (2012): “Building Employee Commitment to Change Across Organizational Levels: The Influence of Hierarchical Distance and Direct Managers’ Transformational Lead,” *Organization Science*, 23, 758–777.
- IDSON, T. L. AND W. Y. OI (1999): “Firm Size and Wages,” in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and D. Card, Elsevier, vol. 3, Part B, chap. 33, 2165–2214.
- IMBENS, G. W. AND J. M. WOOLDRIDGE (2009): “Recent developments in the econometrics of program evaluation,” *Journal of Economic Literature*, 47, 5–86.
- JOHNSON, R. C. (2018): “Measuring Global Value Chains,” *Annual Review of Economics*, 10, 207–236.
- KHANNA, T. AND Y. YAFEH (2007): “Business Groups in Emerging Markets: Paragons or Parasites?” *Journal of Economic Literature*, 45, 331–372.
- KLOSTERHUBER, W., P. LEHNERT, AND S. SETH (2016): “Linked-Employer-Employee-Daten des IAB: LIAB-Querschnittsmodell 2 1993-2014 (LIAB QM2 9314),” *FDZ Datenreport 05/2016*.
- MINCER, J. (1958): “Investment in Human Capital and Personal Income Distribution,” *Journal of Political Economy*, 66, 281–302.

- QIAN, Y. (1994): “Incentives and Loss of Control in an Optimal Hierarchy,” *Review of Economic Studies*, 61, 527–544.
- RAJAN, R. G. AND J. WULF (2006): “The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies,” *The Review of Economics and Statistics*, 88, 759–773.
- ROSENBAUM, P. R. AND D. B. RUBIN (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70, 41–55.
- (1985): “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score,” *The American Statistician*, 39, 33–38.
- TRIGUERO-SÁNCHEZ, R., J. PEÑA VINCES, AND M. SÁNCHEZ-APELLÁNIZ (2013): “Hierarchical Distance as a Moderator of HRM Practices on Organizational Performance,” *International Journal of Manpower*, 34, 794–812.
- UNCTAD (2016): *World Investment Report 2016: Investor Nationality: Policy Challenges*, UNCTAD, Geneva.
- WINTER-EBMER, R. AND J. ZWEIMÜLLER (1999): “Firm-Size Wage Differentials in Switzerland: Evidence from Job-Changers,” *The American Economic Review*, 89, 89–93.