FIRM ORGANIZATION WITH MULTIPLE ESTABLISHMENTS*

Anna Gumpert**

Henrike Steimer

Manfred Antoni

Abstract

We show theoretically and empirically that the managerial organization of multi-establishment firms is interdependent across establishments. To derive our result, we study the effect of geographic frictions on firm organization. In our model, we assume that a CEO's time is a resource in limited supply, shared across headquarters and establishments. Geographic frictions increase the costs of accessing the CEO. Hiring middle managers at one establishment substitutes for CEO time, which is reallocated across all establishments. Consequently, geographic frictions between the headquarters and *one* establishment affect the organization of *all* establishments of a firm. Our model is consistent with novel facts about multi-establishment firm organization that we document using administrative data from Germany. We exploit the opening of high-speed railway routes to show that not only the establishments directly affected by faster travel times but also the other establishments of the firm adjust their organization. Our findings imply that local conditions propagate across space through firm organization.

JEL codes: D21, D22, D24.

Keywords: firm organization, multi-establishment firm, knowledge hierarchy, geography.

*We are deeply grateful to Lorenzo Caliendo for in-depth comments and to Monika Schnitzer for her invaluable support. We are thankful to the editor Pol Antràs and four anonymous referees, who offered many constructive suggestions that are incorporated in this draft. We thank Pol Antràs, Damian Clarke, Wolfgang Dauth, and Alexandra Roulet for very helpful discussions of the paper, and Andrew Bernard, Florian Englmaier, Robert Gibbons, Andreas Steinmayr Martin Watzinger, Jens Wrona as well as seminar participants at Harvard, LSE, LMU Munich, MIT, Oxford, Regensburg, and the Graduate Institute, Geneva, and participants at the CEPR Organizational Economics Conference, NBER Organizational Economics Meeting, CESifo Area Conference on the Global Economy, Verein für Socialpolitik Annual Conference and Committee for Organizational Economics Meeting, and the European Economic Association Conference, for their comments. Amelie Grosenick and Christopher König provided outstanding research assistance. We thank Stefan Seth for his support with assembling the Social Security data. We thank DB Fernverkehr AG for providing the data on the travel times between cities in the German long distance railway network. The paper was partly written while Anna Gumpert was a visitor at the Yale School of Management (SOM). She thanks the SOM for its hospitality and Lorenzo Caliendo for making the visit possible. The project has benefited from financial support through the German Research Foundation (DFG) under CRC TRR 190 (project number 280092119) and GRK 1928, and through a travel grant of the Fritz Thyssen Foundation. An early version of the paper circulated under the title "Communication and the organization of firms across space."

**LMU Munich, Seminar for Comparative Economics, Akademiestr. 1, 80799 Munich, Germany, CEPR, and CESifo, e-mail: anna.gumpert@econ.lmu.de, phone: +49 89 2189 3232, fax: +49 89 2180 2767.

Word count (including Tables and Figures, excluding Appendix): *15,775.*

I. Introduction

Firm performance depends on firm organization, and firm organization crucially depends on firm characteristics. Most notably, it is well known that large firms have more layers of middle managers than small firms. However, firm size alone fails to capture the complexity of large corporations. Large corporations often operate multiple establishments in different locations. The decision to maintain multiple establishments likely affects the managerial organization, as the establishments share managerial resources of the headquarters, and at the same time often encounter diverse conditions in their respective locations. Yet, the effects of operating multiple establishments on the managerial organization are only poorly understood.

This paper studies the managerial organization of firms with multiple establishments. We show that the managerial organization of an establishment does not only depend on the characteristics and local conditions of the establishment, but also on those of the other establishments of the firm. In multi-establishment firms, organization should therefore not be studied at the establishment level, but by considering the firm as a whole. A key implication of our study is that local economic conditions propagate across space through firm organization, because local conditions affect not only the organization of the local establishment, but also the organization of the headquarters and other establishments of a multi-establishment firm.

Specifically, we study the effect of geographic frictions on the managerial organization of a firm and its establishments. We use new data from administrative sources in Germany to document that distance between the establishments and the headquarters increases the number of managerial layers both at the establishments and the headquarters. We develop a model to show that geographic frictions increase the optimal number of managerial layers of multi-establishment firms. Importantly, the model predicts that geographic frictions between the headquarters and one establishment affect not only the optimal managerial organization of this particular establishment, but also the organization of the headquarters and other potential establishments of a firm. We use our data to show that this prediction is reflected in the organizational response of multi-establishment firms to an exogenous reduction in travel times following the opening of high-speed railway routes.

We motivate our study by documenting three facts that, taken together, suggest that the managerial organization of multi-establishment firms is interdependent across establishments. First, the probability that a firm operates an establishment at a location decreases with distance from the headquarters. Distance also correlates negatively with establishment size (in line with Giroud 2013; Kalnins and Lafontaine 2013).

Second, the number of managerial layers of a multi-establishment firm correlates positively with the distance of its establishments from the headquarters (see Figure I). Quantitatively, doubling the distance is associated with the same increase in number of layers as increasing sales by a third. The correlation is not driven by larger firms investing in more distant locations. Distance correlates positively with the number of managerial layers both at the establishments and the headquarters.

Third, multi-establishment firms typically add or drop managerial layers either at the headquarters or the establishments. Only rarely do they alter the number of layers simultaneously at the headquarters and the establishments. This pattern is independent of the distance of establishments.

We propose a model to understand how geographic frictions affect the optimal managerial organization. We model firms as knowledge hierarchies (Garicano 2000; Caliendo and Rossi-Hansberg 2012). We select this framework because recent evidence suggests that the efficient transfer of intangible inputs such as managerial knowledge is an important motive for integrating multiple establishments (Atalay et al. 2014), and that spatial frictions impede knowledge flows both within and between the sites of an organization (Battiston et al. 2020; Keller and Yeaple 2013). Furthermore, the notion of a layer in this framework is closely in line with our measure of layers in the data. We assume that a firm consists of a headquarters and possibly an additional establishment. The production workers at the headquarters and the establishment share a chief executive officer (CEO), who is located at the headquarters. Production is a problem-solving process. Workers input labor and generate problems that must be solved to produce output. The CEO helps the workers solve the problems that they cannot solve using their knowledge. The firm may choose to hire a layer of local middle managers, who solve some of the problems that would otherwise need to be solved by the CEO, but entail a quasi-fixed cost for the firm.

Helping workers costs CEO time. The driving forces of the model are that the CEO has only one unit of time, and that geographic frictions between the establishment and the headquarters increase the amount of time that the CEO needs to help the workers at the establishment.

Through straining CEO time, geographic frictions reduce the probability that a firm operates an establishment. For the same reason, establishments tend to be smaller than the headquarters.

This result is consistent with the lower investment probability and the lower size of establishments at distant locations documented in Fact 1.

The firm adjusts the establishment's organization in response to more severe geographic frictions so that fewer problems need to be solved by the CEO. In particular, geographic frictions render it desirable to hire middle managers. Given that the CEO is shared between the headquarters and the establishment, the firm additionally adjusts the organization at the headquarters. The model thus explains Fact 2: the number of layers increases with geographic frictions, and the managerial organization responds both at the establishments and the headquarters.

As the middle managers entail a quasi-fixed cost, a firm only hires them if firm size is sufficiently large. Importantly, hiring middle managers at the establishment also increases efficiency at the headquarters (and vice versa). This is because middle managers release CEO time, hence middle managers at the establishment increase the amount of CEO time available for the headquarters and reduce the need to hire middle managers there. This result explains Fact 3: multi-establishment firms do not add layers at the headquarters and the establishments at the same time. Both the successive reorganization and the impact of geographic frictions reflect how multi-establishment firm organization is interdependent across establishments.

In the final part of our paper, we utilize the opening of high-speed railway routes in Germany to study the response of firm organization to exogenous variation in geographic frictions. The routes reduce travel time between establishments and headquarters, providing the quickest mode of travel between locations. We focus on the model prediction that geographic frictions between the headquarters and one establishment have repercussions for the managerial organization of the headquarters and other potential establishments of the firm. Importantly, geographic frictions affect establishment size in the model. Size changes lead to changes in the number of layers. Travel times therefore have an indirect effect through size on the managerial organization, in addition to their direct effect. Only the total—direct and indirect—effect of lower travel times is identified.

We find that establishments that benefit from lower travel times grow faster than those that do not. The number of managerial layers is constant. This is consistent with the direct negative effect of lower travel times on the number of layers and the indirect positive effect through larger size compensating each other. Importantly, we find that lower travel times increase the wages and number of managerial layers at the headquarters. This finding supports the interdependence of the managerial organization predicted by the model. The interdependence goes beyond the headquarters: if a firm has at least one establishment affected and one unaffected by lower travel times, the wages and share of employees in managerial occupations in the unaffected establishment increase faster than in establishments of firms that do not benefit at all from lower travel times.

Through the lens of the model, lower travel times between the headquarters and the establishment affect the managerial organization because they decrease the costs of accessing CEO knowledge at the establishment. In supplementary regressions, we exploit the model's implication that changes in the helping costs have a more pronounced impact in sectors with a less predictable production process to explore this channel. We construct a sector-level measure of predictability using survey data on the tasks and workplace environment of employees. We find that the estimated effects are driven by establishments and headquarters in sectors with below-median predictability of the production process. This evidence supports the mechanism proposed by the model. In addition, we find that the education and experience of employees change concomitantly with wages, which also supports a knowledge hierarchy model.

Our paper contributes to several strands of the literature. To develop our model, we build on the literature of firms as knowledge hierarchies (for an overview, see Garicano and Rossi-Hansberg 2015). Our paper is closest to that of Antràs et al. (2008), which shows that middle managers facilitate the transmission of knowledge in the context of offshoring. Our model goes beyond their theory by incorporating simultaneous production at the headquarters and the establishment of a firm. This enables us to study the effect of local shocks on the organization of the local and non-local units of a firm. The broader literature focuses on firm size as a determinant of organization (Caliendo and Rossi-Hansberg 2012, henceforth CRH; Caliendo et al. 2015, 2020; Friedrich 2020).[1] The possibility of multi-establishment production is largely neglected, although multi-establishment firms account for a substantial share of aggregate employment in developed economies.[2]

Our main result is that the managerial organization of multi-establishment firms is interdepen-

[1]Using a knowledge hierarchy model, Mariscal (2018) shows that the impact of new information technologies on firm organization explains the decline of the US labor share. Spanos (2019) shows that firm organization explains part of the productivity differences across locations. In the empirical literature on firm hierarchies, Rajan and Wulf (2006) document the flattening of corporate hierarchies over time. Guadalupe and Wulf (2010) examine the impact of competition on corporate hierarchies. Sforza (2020) compares the organizational responses to a credit supply and a trade shock.

[2]Gumpert (2018) develops a knowledge hierarchy model with multiple establishments, but a fixed number of layers. Crèmer et al. (2007) study firm language in a setting with multiple divisions. McElheran (2014) presents facts about the allocation of decision-making authority in multi-establishment firms based on team-theoretic considerations.

dent across establishments. The key driver of this result is that the establishments share the CEO. Hence, our result is more general than our specific model of firm organization, in line with prior findings on structural similarities of different hierarchy models (Chen 2017; Chen and Suen 2019).

The interdependence of establishment organization is particularly relevant for recent literature documenting how multi-establishment firms propagate local shocks through their internal networks (Giroud and Mueller 2019; Seetharam 2018). This body of research discusses managerial and financial constraints as potential drivers of the empirical findings. However, although CEOs are considered decisive for firm performance (Bertrand 2009), managerial constraints have received very little systematic attention. Our contribution is to provide both a formal analysis and empirical evidence regarding the role of managerial constraints for multi-establishment firm organization.

Our empirical strategy builds on literature using the opening of high-speed railway routes to identify the impact of geographic frictions on firms (e.g., Bernard et al. 2019). Our approach is particularly close to Charnoz et al. (2018), who study the effect of new routes on the functional specialization and hierarchical organization of business groups. Their results are consistent with the predictions of our model, although they use different outcome variables to capture both aspects of firm organization. Our contribution is to provide a unified theoretical and empirical analysis. Our outcome variables neatly map firm organization in our model. Our model explains why the impact of geographic frictions goes beyond a particular establishment and cleanly disentangles the direct effects of geographic frictions on organization and the indirect effects through size.

Our paper also relates to the literature on multinational firms. Keller and Yeaple (2013) back out the costs to transfer knowledge across space from multinational operations. Their results support our assumption that geographic frictions hamper access to headquarter knowledge. In the broader literature, headquarter inputs are often considered public goods within the firm (e.g., Helpman et al. 2004; Irarrazabal et al. 2013; Antràs and Yeaple 2014, for a survey). Our results caution that this assumption may apply to patents or trademarks, but not necessarily to managerial inputs.

Finally, our paper offers a novel perspective on the recent management literature. Bloom et al. (2019) document that half of the total variation in management practices between US establishments owes to variations between establishments within the same firm. Implementing managerial practices requires managerial time. The heterogeneity of management practices may reflect asymmetries in the number of layers and the amount of CEO time allocated to an establishment.

The paper is structured as follows. Section II describes the data. Section III presents the facts on multi-establishment firm organization. Section IV develops the model. Section V presents the evidence from the opening of high-speed railway routes. The final section concludes.

II. Data

II.A. Data sources

We use a linked firm-establishment-employee data set for Germany that is uniquely suited to the study of multi-establishment firms. The data contain information on the sales and legal form of firms, as well as the county and the sector of their establishments. For each establishment, we observe all employees subject to social security contributions on 30 June, and their occupation and wage. The data cover firms in all sectors during the period 2000–2012. Each employee, establishment and firm has a unique identifier that makes it possible to follow them over time.

We assemble the data set from two sources. The universe of social security records provides the data on employees and establishments. The Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research makes these data available for research. We use the employee history, the Establishment History Panel and the extension files on entries and exits of establishments. The Orbis database of Bureau van Dijk contains the balance sheet information of firms. We use a linkage table between the social security records and the Orbis database. The headquarters (HQ) of a firm is identified as the establishment with the same zip code or locality as the firm.[3] Appendix A.1 contains details on the components of our data set and the record linkage procedure.

The data set is an unbalanced panel. We exclude the year 2011 due to changes in the occupational classification in that year (see Appendix A.2). Our main analyses use the 2000–2010 panel. We use the year 2012 for cross-sectional analyses, because it contains the maximum number of establishments, exhibits relatively few missing values for sales, and uses the new, finer occupational classification. Consistent with the literature, we restrict our sample to full-time employees (e.g., Card et al. 2013). We focus on firms with at least 10 employees in all years.

[3]The social security data contain the address of an establishment. We are not allowed to use the address for our empirical analyses due to data confidentiality.

Multi-establishment (ME) firms comprise the headquarter establishment and at least one other establishment. For clarity, we use the term "headquarters" for the former and "establishment" to denote the latter. Single-establishment (SE) firms only consist of the headquarters.

II.B. Measures for managerial organization

We use the occupation of the employees to construct three measures of the managerial organization of firms. Our preferred measure is the number of managerial layers. We assign employees to four layers (following Caliendo et al. 2015, see Appendix A.3 for details):

Level	Designation	Occupations
3	CEO	CEOs, managing directors
2	Middle managers	Senior experts, middle managers
1	Supervisors	Supervisors, engineers, technicians, professionals
0	Production workers	Clerks, operators, production workers

We treat the layer at the lowest level in the firm as non-managerial and count the number of layers above the lowest layer per firm.

Alternatively, we use shares of managerial occupations in the wage sum. The establishments report the occupations of employees in the social security data. In ME firms, establishments may assign different occupations to similar employees. Cross-checking the results regarding the number of layers with the managerial share ensures that our results are robust to this possibility. We determine managerial occupations in two ways. On the one hand, we use the assignment of employees to layers and treat all employees above the lowest level as managerial. On the other hand, we use the Blossfeld (1983, 1987) occupational categories, which build on research from sociology and are part of the Establishment History Panel.

Appendix A.4 illustrates the plausibility of the assignment of employees to layers. We replicate Caliendo et al. (2015) and show how the tasks of employees systematically differ between layers in ways that plausibly reflect the different roles of employees within firms using survey data.

II.C. Descriptive statistics

Table I provides descriptive statistics of the 2012 cross-section (see Appendix A.5.1 for the 2000-2010 data). Our sample comprises 109,500 firms. We only observe sales for the larger firms owing to missing values in the Orbis data. The firms consist of 144,500 establishments (including headquarters) and employ 6.4 million individuals. Nine percent of firms are ME firms. They make up a disproportionate share of establishments and employment: 31 percent of establishments belong to them, and 34 percent of employees work for them. This pattern is similar across sectors (see Table A.16). ME firms are substantially larger than SE firms in terms of employment and sales. ME firms also have a higher number of managerial layers and higher managerial shares. Appendix A.5.2 shows that this difference does not only reflect differences in firm size.

Table II illustrates the complexity of ME firms. On average, ME firms have five establishments (including headquarters). Half of them have two, and the largest five percent have 10 or more establishments. The establishments tend to be geographically dispersed. At the top of the distribution, the air-line distance between headquarters and establishments exceeds 540 km, about two thirds of the maximum possible distance within Germany. Headquarters are substantially larger than establishments. Management is concentrated in the headquarters that have a higher number of managerial layers and higher managerial shares than establishments. Appendix A.5.3 describes the organization of headquarters and establishments in detail.

III. Facts

III.A. Distance to headquarters decreases location probability

Table III describes the geographic organization of ME firms. Columns 1 to 3 show that firms are less likely to locate an establishment in a county that is distant from their headquarters. According to columns 4 to 6, establishment size also decreases with distance. Larger market potential increases location probability and establishment size. Higher wages and land prices in the county relative to the headquarters are negatively associated with location probability. Although higher wages also relate negatively to establishment size, higher land prices relate positively.

The results are consistent with a negative impact of geographic frictions between the headquarters and an establishment on establishment performance. The effects of market potential and wages indicate market-seeking and cost-cutting motives for having establishments. The different effects of land prices on location decision and size are in line with the cost of land being a fixed cost, so it is worth maintaining only larger establishments at locations with higher land prices.

Fact 1 summarizes our findings:

Fact 1. *Distance of a county from the headquarters is negatively related to the probability that a ME firm locates an establishment there as well as the size of the establishment conditional on location.*

Appendix B.1 presents the results graphically and in the 2000–2010 panel.

III.B. Distance to headquarters increases the number of layers

Figure I in the introduction shows that the number of managerial layers relates positively to the distance between the headquarters and the establishments. Table IV documents that the relationship is robust to controlling for size to capture the positive effect of size on the number of layers (e.g., Caliendo et al. 2015) and the possibility of larger firms investing in more distant locations. We estimate Poisson regressions:

$$\#\,\text{managerial layers}_i = \exp\left(\beta_0 + \beta_1 \text{geographic frictions}_i + \beta_2 \text{size}_i + \alpha_l + \alpha_n + \alpha_s\right)$$

i refers to the firm, l to its legal form, n to the county of the headquarters, s to the headquarter sector, and α denotes fixed effects. To account for the fractional nature of the managerial share, we follow Papke and Wooldridge (1996) and estimate a generalized linear model.

We approximate geographic frictions with the maximum distance of establishments to the headquarters or the minimum area spanned by the establishments and the headquarters. The distance is defined for all ME firms, whereas the area is only defined for firms active in at least three counties. We use sales and the number of non-managerial employees as measures of firm size.

The regression results show that both distance and area relate positively with the number of managerial layers in a firm. According to column 3, doubling the maximum distance of establishments to the headquarters is associated with the same increase in the number of layers as 46 percent more non-managerial employees. Moving from the lower to the upper quartile of the distribution of distance and the number of non-managerial employees is associated with 0.2 and 0.6 more layers, respectively. Taken together, this accounts for about half of the interquartile range of the number of layers. The managerial share also relates positively to the distance and the area.

The firm-level results may disguise different responses of headquarter and establishment organization. The managerial organization of the establishments does not copy the headquarters':

71 percent of establishments have fewer managerial layers than the headquarters. We therefore complement the firm-level results with establishment and headquarter level analyses. Figure II shows that the number of managerial layers in both the headquarters and the establishments relates positively to distance. According to Table V, doubling the (maximum) distance is associated with the same increase in the number of layers as 22 percent more non-managerial employees in the establishments and 35 percent more non-managerial employees in the headquarters.

Fact 2 summarizes our findings:

Fact 2. *The number of managerial layers of ME firms correlates positively with the distance between headquarters and establishments and the area they span, conditional on firm characteristics. The number of managerial layers of both the establishments and the headquarters increases with distance.*

Appendix B.2 documents that the results are robust to modifications of the main variables, to alternative econometric specifications, in the 2000–2010 panel, and in sample splits.

III.C. Reorganization of headquarters or establishments

To complement the cross-sectional evidence on the managerial organization, we study the reorganization dynamics of firms over time. The upper panel of Table VI displays the share of ME firms that transition from a number of managerial layers in year t to a potentially different number of layers in year $t + 1$. At least 80 percent of firms keep the number of layers constant. If they alter the number of layers, firms usually add or drop one layer. These dynamics are similar to those of French and Danish firms (Caliendo et al., 2015b, Friedrich, 2020) and SE firms (Table B.18).

The lower panel displays the reorganization dynamics at the level of the headquarters and establishments. We count the maximum number of layers at the establishments to account for the potentially different number of establishments across firms. Over time, the managerial organization at the unit level is less stable than the managerial organization at the firm level as reflected by less mass on the diagonal of the lower panel than on the diagonal of the upper panel. Notably, if ME firms change their organization, they typically add or drop layers at either the headquarters or the establishment(s), but not both. For example, among ME firms with two layers at the headquarters and the establishments, nine percent add one layer at the headquarters and 10 percent drop one layer at the establishments. Only two percent choose a lower or higher number of layers at both.

Among the firms that change the number of layers, 49 percent change it only at the headquarters, 42 percent change it only at the establishments, and just nine percent change it at both.[4] Figure III illustrates that changes of the number of layers of the headquarters or establishments are not necessarily visible at the firm level. 34 percent of changes of the number of layers at the headquarters are reflected in a lower firm-level number of layers, and 37 percent in a higher firm-level number of layers. The remaining 29 percent are not visible at the firm level (Figure IIIa). 65 percent of changes of the number of layers at the establishments do not change the firm-level number of layers (Figure IIIb). If firms change the number of layers simultaneously at the headquarters and the establishments, 40 percent of the changes are not visible at the firm level (Figure IIIc).

Fact 3 summarizes our finding.

Fact 3. *ME firms that reorganize typically add or drop layers either at the headquarters or at the establishments.*

Appendix B.3 documents that changes in the number of layers are related to changes in firm size (consistent with Caliendo et al. 2015; Friedrich 2020), and shows that our results are robust to different ways of counting managerial layers, longer time lags, and in sample splits.

IV. Model

IV.A. Set-up

We use Facts 1 to 3 to inform a model in which firms endogenously choose whether to operate an establishment and the managerial organization. We consider an economy with two locations, $j = \{0, 1\}$. N_j agents each supply one unit of time to the labor market in location j. Agents work in two sectors. In the differentiated-goods sector, each firm i produces one product. The homogeneous-good sector produces a non-tradeable good under perfect competition using a constant-returns-to-scale technology that may differ between locations. The agents are mobile between sectors, but immobile between locations, so wages w_j may differ. The agents consume the homogeneous good and the differentiated products.

[4] All figures refer to all firms, i.e., they include firms that have a higher number of layers at the establishment than at the headquarters and are excluded from Table VI.

Production. Production in the differentiated-goods sector is a problem-solving process based on labor and knowledge (Garicano 2000; CRH). One unit of labor generates a unit mass of problems that have to be solved using knowledge to produce output. Mathematically, knowledge is an interval ranging from zero to an upper bound. We denote the length of the interval by z. A problem is solved if it is realized within the interval. The problems follow a distribution with the exponential density $f(z) = \lambda e^{-\lambda z}$, where $z \in [0, \infty)$ refers to the domain of possible problems and λ denotes the predictability of the production process. A higher value of λ implies that a given amount of knowledge solves more problems. Combining n units of labor and knowledge $\bar{z}$ yields $q = n\left(1 - e^{-\lambda \bar{z}}\right)$ units of output, where $1 - e^{-\lambda \bar{z}}$ is the value of the cumulative distribution function.

A firm hires agents for production. The firm's employees supply labor by spending their time generating problems. To supply knowledge, employees must learn. They spend $w_j cz$ to learn knowledge z, where c denotes the learning cost. The firm remunerates the employees for their time and learning expenses, so they receive remuneration $w_j(1 + cz)$ (as in CRH).

The employees of the firm can share problems among themselves, and hence can leverage differences in knowledge. This is costly: an employee in location j spends θ_{kj} units of time helping an employee in location k. Helping is more costly across than within locations: $1 > \theta_{10} \geq \theta_{00} > 0$. The helping costs are symmetric: $\theta_{10} = \theta_{01}, \theta_{11} = \theta_{00}$. If an employee does not know how to solve a problem, he cannot tell who knows, but must find a competent fellow employee.

Organization. Firms organize their employees in hierarchical layers. We call the employees at the lowest layer $\ell = 0$ production workers. They supply labor and solve the problems realized in their knowledge interval. We call the employees at the higher layers $\ell \geq 1$ managers. They supply only knowledge and spend their time helping the employees at the next lowest layer. The CEO constitutes the highest managerial layer. All firms consist at least of production workers and a CEO; they may also have one or more layers of middle managers. The knowledge levels of the employees are overlapping, so employees at layer ℓ know the knowledge of employees at layer $\ell - 1$ and more.[5] The CEO is the most knowledgeable employee of the firm. A crucial assumption is that each firm has exactly one CEO, who is thus a resource in limited supply for a firm.

The helping costs θ_{jk}, learning costs c, and the predictability of the production process λ are ex-

[5] Assuming overlapping knowledge levels simplifies the analysis, but does not drive the results (see Appendix C.1.2).

ogenous parameters, with values restricted by Assumption 1 (Appendix C.1.1). The homogeneous-good sector pins down wages w_j. To simplify the exposition, sections IV.B and IV.C examine the organization of a firm in location 0 taking output as given. Section IV.D endogenizes output.

IV.B. Single-establishment firm organization

We first determine the optimal organization of a SE firm as a benchmark for the analysis of ME firm organization. The organization consists of the number of below-CEO layers of middle managers L, the number $n_{0,L}^{\ell}$ and knowledge $z_{0,L}^{\ell}$ of employees per layer ℓ, and the knowledge of the CEO $\bar{z}_{0,L}$. The indeces $0, L$ refer to the location of the firm and the number of below-CEO layers, reflecting that these variables affect the other choices. We focus on the decision to hire one layer of middle managers and study the decision to hire additional layers in Appendix C.2.1.

The optimal organization yields minimal production costs:

$$C(\tilde{q}) = \min_{L} \tilde{C}_{0,L}(\tilde{q}) \tag{1}$$

$$\text{where} \quad \tilde{C}_{0,L}(\tilde{q}) = \min_{\{n_{0,L}^{\ell}, z_{0,L}^{\ell}\}_{\ell=0}^{L}, \bar{z}_{0,L} \geq 0} \left[\sum_{\ell=0}^{L} n_{0,L}^{\ell} w_0 \left(1 + c z_{0,L}^{\ell}\right) \right] + w_0 \left(1 + c \bar{z}_{0,L}\right) \tag{2}$$

$$\text{s.t.} \quad n_{0,L}^{0} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right) \geq \tilde{q} \tag{3}$$

$$1 \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{L}} \tag{4}$$

$$n_{0,L}^{\ell} \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{\ell-1}} \quad \text{for } \ell > 0 \tag{5}$$

$$\bar{z}_{0,L} \geq z_{0,L}^{L}, \quad z_{0,L}^{\ell} \geq z_{0,L}^{\ell-1} \quad \text{for } \ell > 0 \tag{6}$$

The production costs consist of the costs for the employees and the CEO. Constraint (3) specifies that the number of production workers and CEO knowledge must suffice to produce output. Constraints (4) and (5) reflect that the CEO and the middle managers have only limited time to help production workers solve problems. Knowledge levels are overlapping (constraint 6).

Appendix C.2.2 contains the Lagrangian equation and the first order conditions. Two multipliers from the Lagrangian equation help characterize the organization. The multiplier for constraint (3), $\xi_{0,L}$, denotes the marginal production costs. The multiplier for constraint (4), $\varphi_{0,L}$, denotes the marginal benefit of CEO time that reflects how costly the CEO time constraint is for the firm.

CEO knowledge is optimal if the marginal increase of CEO remuneration equals the marginal decrease of production costs:

$$w_0 c = \xi_{0,L} n^0_{0,L} \lambda e^{-\lambda \bar{z}_{0,L}} \tag{7}$$

Constraints (3)-(5) determine the number of production workers, the knowledge at the highest below-CEO layer, and the number of middle managers (if any). If the firm hires middle managers, the knowledge of the production workers is a function of the managerial knowledge:

$$e^{\lambda z^0_{0,L}} = \left(1 + c z^1_{0,L}\right) \frac{\lambda \theta_{00}}{c}. \tag{8}$$

The marginal production costs $\xi_{0,L}$ and the marginal benefit of CEO time $\varphi_{0,L}$ are:

$$\xi_{0,L} = \frac{w_0 \left(1 + c z^0_{0,L} + \frac{c}{\lambda} + \mathbb{1}(L=1) \frac{c}{\lambda} \theta_{00} e^{-\lambda z^0_{0,L}}\right)}{1 - e^{-\lambda \bar{z}_{0,L}}},$$
$$\varphi_{0,1} = \frac{w_0 c}{\lambda} e^{\lambda \left(z^1_{0,L} - z^0_{0,L}\right)} \quad \text{for } L = 1, \quad \varphi_{0,0} = \frac{w_0 c}{\lambda \theta_{00}} e^{\lambda z^0_{0,0}} \quad \text{for } L = 0.$$

Understanding how output $\tilde{q}$ affects firm choices is useful for the analysis of ME firms.

Proposition 1. *Given the number of below-CEO managerial layers L of the firm,*

a) the knowledge of the CEO $\bar{z}_{0,L}$, the number $n^\ell_{0,L}$ and the knowledge $z^\ell_{0,L}$ of the employees at all below-CEO layers $\ell \leq L$, the managerial span of control $n^{\ell-1}_{0,L}/n^\ell_{0,L}$ at all managerial layers $1 \leq \ell \leq L+1$, and the marginal benefit of CEO time $\varphi_{0,L}$ increase with output $\tilde{q}$.

b) The cost function strictly increases with output $\tilde{q}$. The marginal costs increase with output for $\tilde{q} \geq \hat{q}^L$, with $\hat{q}^0 = 0$. The average cost function is U-shaped. It reaches a minimum at $\tilde{q}^{*L}$ where it crosses the marginal cost function, and converges to infinity for $\tilde{q} \to 0$ and $\tilde{q} \to \infty$.

Proof. See Appendix C.2.3. □

Our results align with those in CRH for non-overlapping knowledge. We consider an additional outcome, the marginal benefit of CEO time. It increases with output, because higher output makes it more beneficial to increase CEO time and avoid the increase in below-CEO knowledge.

Regarding the decision to hire middle managers, the firm faces a trade-off. On the one hand, middle managers entail a quasi-fixed cost, because they are remunerated but do not generate problems. On the other hand, middle managers reduce the number of problems sent to the CEO and thus allow decreasing the knowledge of the production workers and the marginal production costs. Consequently, hiring middle managers is only worthwhile if the firm is sufficiently large.

Figure IVa illustrates the choice of hiring middle managers. The minimum efficient scale $\tilde{q}^{*L}$ increases with the number of below-CEO layers, reflecting the higher quasi-fixed costs. The cost function becomes flatter with the number of layers, because the marginal production costs increase less strongly with output. The firm adds a layer at the crossing $\tilde{q}^{0\to 1}$ (see Appendix C.2.4).

IV.C. Multi-establishment firm organization

The firm may maintain an establishment at location 1 to exploit wage differences or improve access the local output market. To capture market access motives, we assume that the firm incurs iceberg-type transport costs $\tau \geq 1$ to ship output produced in one location to the other location.[6] We take as given the potentially different amounts of output $\tilde{q}_j$ that the firm supplies at each location.

The CEO is located at the headquarters in location 0. The firm chooses whether to produce only in the headquarters, the establishment or both, as well as the number of below-CEO layers of middle managers L_j per location. We use the term "organizational structure" and the variable ω to denote the combination of the below-CEO layers (L_0, L_1). All other endogenous variables depend on the location and the organizational structure, so we index them by j, ω.

We split the optimization problem into three steps. First, the firm chooses the organizational structure ω, similarly to choosing the number of layers in section IV.B:[7]

$$C\left(\{\tilde{q}_j\}_{j=0}^{1}\right) = \min_{\omega} \tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^{1}\right) \tag{9}$$

Second, the firm chooses how much output $q_{j,\omega}$ and which share $s_{j,\omega}$ of CEO time to allocate to the headquarters and the establishment as well as CEO knowledge $\bar{z}_{0,\omega}$:

[6] I.e., $\tau \geq 1$ units of a good need to be shipped for one unit to arrive at destination.

[7] As in section IV.B, we focus on the decision to hire the first managerial layer and study hiring additional layers in Appendix C.3.1.

$$(10)\quad \tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^{1}\right) = \min_{\{q_{j,\omega},s_{j,\omega}\}_{j=0}^{1},\bar{z}_{0,\omega}\geq 0} \sum_{j=0}^{1} C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0\left(1 + c\bar{z}_{0,\omega}\right)$$

$$(11)\qquad \text{s.t.}\quad s_{0,\omega} + s_{1,\omega} \leq 1$$

$$(12)\qquad q_{0,\omega} + q_{1,\omega} \geq \tilde{q}_0 + \tilde{q}_1$$

$$(13)\qquad q_{j,\omega} \geq \tilde{q}_j + \tau(\tilde{q}_k - q_{k,\omega}) \quad \text{if } q_{k,\omega} \leq \tilde{q}_k,\ k \neq j$$

The costs consist of the costs per location and the remuneration of the CEO time that is not used in production. Equation (11) reflects the CEO's time constraint. Equation (12) states that total production has to cover total output. Local production may be lower than local output. Equation (13) states that if production is lower than output at one location, production at the other location has to compensate the shortfall plus transport costs.

Third, the firm chooses the number $n_{j,\omega}^{\ell}$ and knowledge $z_{j,\omega}^{\ell}$ of the employees in each layer ℓ.

$$(14)\quad C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) \begin{cases} \overset{q_{j,\omega}>0}{=} \min\limits_{\{n_{j,\omega}^{\ell},z_{j,\omega}^{\ell}\}_{\ell=0}^{L_j}\geq 0} \left[\sum_{\ell=0}^{L_j} n_{j,\omega}^{\ell} w_j\left(1 + c z_{j,\omega}^{\ell}\right)\right] + s_{j,\omega} w_0\left(1 + c\bar{z}_{0,\omega}\right) \\ \overset{q_{j,\omega}=0}{=} 0 \end{cases}$$

$$(15)\qquad \text{s.t.}\quad n_{j,\omega}^{0}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right) \geq q_{j,\omega}$$

$$(16)\qquad s_{j,\omega} \geq n_{j,\omega}^{0}\theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}}$$

$$(17)\qquad n_{j,\omega}^{\ell} \geq n_{j,\omega}^{0}\theta_{jj} e^{-\lambda z_{j,\omega}^{\ell-1}} \quad \text{for } \ell > 0$$

$$(18)\qquad \bar{z}_{0,\omega} \geq z_{j,\omega}^{L_j}, \quad z_{j,\omega}^{\ell} \geq z_{j,\omega}^{\ell-1} \quad \text{for } \ell > 0$$

The production costs consist of the below-CEO personnel costs and the remuneration for the CEO time allocated to the location. The constraints (15)-(18) are analogous to the constraints (3)-(6).

We solve the problem backwards. We determine the number and knowledge of the employees per layer, taking as given the firm level choices and the organizational structure. We then solve for the firm level choices given the organizational structure, which we determine last. Appendix C.3.2 contains the Lagrangian equations and the first order conditions.

Establishment-level choices. Constraints (15)-(17) determine the number of production workers, the knowledge level of the highest below-CEO layer, and the number of middle managers. As in a SE firm, the knowledge of the production workers is a function of the managerial knowledge. The formal expressions are variants of those in section IV.B, so we state them in Appendix C.3.2. The Lagrangian multipliers $\xi_{j,\omega}$ and $\varphi_{j,\omega}$ denote the marginal production costs and the marginal benefit of CEO time at location j.

Firm-level choices. The firm uses the full unit of CEO time and produces only the given output, i.e., the constraints (11), and (12) or (13) are binding. The firm balances the marginal benefit and marginal cost of CEO knowledge, as in section IV.B:

$$w_0 c = \sum_{j=0}^{1} \xi_{j,\omega} n^0_{j,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}}. \tag{19}$$

Proposition 2 shows how the transport costs affect ME firm organization.

Proposition 2. *Suppose the firm produces in the headquarters and the establishment. The firm allocates CEO time to equalize the marginal benefit of CEO time across locations. Formally, in optimum:*

$$\varphi_{0,\omega} = \varphi_{1,\omega}. \tag{20}$$

The firm either sets the production quantities equal to local output or chooses them to equalize the marginal production costs adjusted by the transport costs across locations. Formally, in optimum,

$$\xi_{0,\omega} \leq \tau \xi_{1,\omega} \wedge \xi_{1,\omega} \leq \tau \xi_{0,\omega} \qquad \textit{if } q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0, \tag{21}$$

$$\tau \xi_{0,\omega} = \xi_{1,\omega} \qquad \textit{if } \tilde{q}_1 > q_{1,\omega} \wedge q_{0,\omega} = \tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega}), \ \textit{and} \tag{22}$$

$$\xi_{0,\omega} = \tau \xi_{1,\omega} \qquad \textit{if } \tilde{q}_0 > q_{0,\omega} \wedge q_{1,\omega} = \tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega}). \tag{23}$$

In the special case of no transport frictions, $\tau = 1$, the firm chooses the production quantities to equalize the marginal production costs across locations:

$$\xi_{0,\omega} = \xi_{1,\omega} \quad \textit{if } \tau = 1, \ \textit{i.e.,} \ q_{1,\omega} = \tilde{q}_0 + \tilde{q}_1 - q_{0,\omega}. \tag{24}$$

Proof. See Appendix C.3.3. □

The firm can flexibly allocate CEO time, so it reallocates CEO time until its marginal benefit is equal across locations.[8] The transport costs limit the flexibility of the allocation of production. The firm has three options: produce output locally, ship it from the other location, or do both. The firm produces output locally if local production has lower marginal costs than shipping output from the other location (equation 21). If local production and shipping output from the other location have the same marginal costs, the firm produces part of the output locally and ships part of it from the other location (equations 22, 23). In the special case of no transport frictions, the firm reallocates quantities until the marginal costs at the headquarters and the establishment are equal. Finally, if shipping output from the headquarters is cheaper than production at the establishment (or vice versa), the firm produces total output in the headquarters (establishment).

Comparative statics. To derive the optimal organizational structure ω, it is useful to understand how choices depend on the output $\tilde{q}_j$ and the helping costs θ_{10}. The comparative statics depend on which of equations (21)-(23) holds. Parameter changes easily lead to a violation of equations (22) and (23). We therefore assume that equation (21) holds. Appendix C.3.8 contains the results for the other cases (including equation 24).

Proposition 3. *Suppose the firm produces in the headquarters and the establishment. Suppose that the firm incurs transport costs $\tau > 1$ to ship output between locations, and that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$ increases with output $\tilde{q}_j$. Higher output $\tilde{q}_j$ increases the number of production workers $n^0_{j,\omega}$ and the share of CEO time $s_{j,\omega}$ at location j and decreases the number of workers $n^0_{k,\omega}$ and the share of CEO time $s_{k,\omega}$ at location $k \neq j$.

b) The knowledge of the employees at all below-CEO layers $z^\ell_{k,\omega}$, $\ell \leq L_k$, $k = 0, 1$, the below-CEO managerial span of control $n^{\ell-1}_{k,\omega}/n^\ell_{k,\omega}$, $1 \leq \ell \leq L_k$, and the marginal benefit of CEO time $\varphi_{0,\omega}$ increase with local output $\tilde{q}_j$ if the CEO spends a sufficient share of time on location j. The marginal production costs $\xi_{k,\omega}$ increase with output $\tilde{q}_j$ if CEO knowledge is sufficiently high.

[8] As $\varphi_{0,\omega} = \varphi_{1,\omega}$, we state results only for $\varphi_{0,\omega}$ in the following.

Proof. See Appendix C.3.4. $\square$

Changes of output $\tilde{q}_j$ have a similar effect on endogenous outcomes at location j as in a SE firm. They also affect outcomes at the other location, where the firm hires fewer production workers due to the common CEO and higher CEO knowledge. The CEO allocates his time accordingly. If the CEO spends a sufficiently high share of time at location j, the increase in the number of production workers and thus problems there outweighs the decrease at the other location. Due to the CEO time constraint, below-CEO knowledge levels increase at both locations. Correspondingly, the marginal benefit of CEO time and the below-CEO managerial span of control rise. The effect on the CEO span of control is ambiguous: the higher number of production workers at location j increases the CEO span of control, but the lower number of production workers at the other location and the higher below-CEO knowledge may outweigh this effect. The marginal production costs increase with output if the firm is large enough, as reflected by sufficiently high CEO knowledge.

Proposition 4. *Suppose the firm produces in the headquarters and the establishment. Suppose that the firm incurs transport costs $\tau > 1$ to ship output between locations, that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$, and that the helping costs are higher across than within locations, $\theta_{10} > \theta_{00}$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$, the knowledge of the employees at all below-CEO layers $z^{\ell}_{1,\omega}$, $\ell \leq L_1$, the managerial span of control $n^{\ell-1}_{1,\omega}/n^{\ell}_{1,\omega}$, $1 \leq \ell \leq L_1$, and the marginal production costs $\xi_{1,\omega}$ at the establishment increase with the helping costs θ_{10}. The total number of production workers $\sum_{j=0}^{1} n^0_{j,\omega}$ as well as the number of production workers $n^0_{1,\omega}$ at the establishment decrease. The share of CEO time $s_{1,\omega}$ decreases if $\exists j$ s.t. $L_j > 0$; it is constant otherwise.

b) The knowledge of the employees at all below-CEO layers $z^{\ell}_{0,\omega}$, $\ell \leq L_0$, the managerial span of control $n^{\ell-1}_{0,\omega}/n^{\ell}_{0,\omega}$, $1 \leq \ell \leq L_0$, the number of production workers $n^0_{0,\omega}$, and the marginal production costs $\xi_{0,\omega}$ at the headquarters as well as the marginal benefit of CEO time $\varphi_{0,\omega}$ decrease with the helping costs θ_{10}. The CEO span of control $\sum_{j=0}^{1} n^{L_j}_{j,\omega}$ decreases with the helping costs θ_{10} if $\exists j$ s.t. $L_j = 0$ or the ratio w_0/w_1 is sufficiently high. The headquarter share of CEO time $s_{0,\omega}$ increases if $\exists j$ s.t. $L_j > 0$; it is constant otherwise.

Proof. See Appendix C.3.5. $\square$

Higher helping costs θ_{10} make it more costly to generate problems at the establishment, because they may have to be sent to the CEO. The firm therefore adjusts establishment organization in order to generate fewer problems and send them to the CEO less frequently. To achieve the former, the firm decreases the number of production workers and increases CEO knowledge. To achieve the latter, the firm increases the knowledge of the employees and thus the below-CEO managerial span of control in the establishment. This allows the CEO to reallocate his time from the establishment to the headquarters, but increases the marginal production costs at the establishment.

The reorganization of the establishment has repercussions for the headquarters. Higher CEO knowledge decreases the number of production workers at the headquarters and thus the total number. Fewer problems are generated, so the knowledge of the employees at the below-CEO layers decreases, as do the below-CEO managerial span of control and the marginal production costs. The CEO span of control decreases due to the lower total number of production workers and the higher knowledge at the establishment, two factors that the lower knowledge at the headquarters does not outweigh. Correspondingly, the marginal benefit of CEO time decreases.

The key implication of Propositions 3 and 4 is that the organization of the ME firm is interdependent across the headquarters and the establishment. Changes in output or helping costs at the establishment result in organizational adjustments at the establishment and the headquarters owing to the shared CEO.

Organizational structure. To render transparent the distinct effects of ME production and location characteristics on organizational structure, we first consider the special case of ME production in two identical locations without transport costs or helping cost frictions. We next add transport costs and helping costs frictions. Finally, we study differences in wages and local output.

Special case: identical location characteristics, no geographic frictions, $w_1 = w_0$, $\theta_{10} = \theta_{00}$, $\tau = 1$. Without transport costs, only total output $\tilde{q}$ matters for the managerial organization. The marginal production costs are equal across locations. This holds mechanically if the number of below-CEO managerial layers is equal. In this case, ME production is equivalent to SE production. If the numbers of below-CEO managerial layers differ, the firm effectively produces with two distinct production technologies, albeit with the same marginal costs. Section IV.B shows that the efficiency of a certain number of layers depends on output for a SE firm. The ME firm can choose the optimal combination of layers for its output by allocating CEO time and production quantities. This affects the choice of optimal organizational structure.

Proposition 5. *Suppose that wages are equal, $w_0 = w_1$, and that there are no transport costs or helping cost frictions, $\tau = 1, \theta_{00} = \theta_{10}$. Let "$(L_0, L_1)$-organization" denote the organizational structure of a ME firm with L_0 below-CEO layers at the headquarters and L_1 below-CEO layers at the establishment.*

a) The average cost function of the (L_0, L_0)-organization is U-shaped in output and reaches a minimum at $\tilde{q}^{*(L_0,L_0)}$.

b) The average cost functions of the (L_0, L_0+1)-organization and the (L_0+1, L_0)-organization coincide. The average cost of the (L_0, L_0+1)-organization is equal to the average cost of the (L_0, L_0)-organization at $\tilde{q}^{*(L_0,L_0)}$, and decreases with output for $\tilde{q} \in (\tilde{q}^{*(L_0,L_0)}, \tilde{q}^{*(L_0+1,L_0+1)})$.

c) The average cost function of the (L_0, L_0)-organization crosses the average cost function of the (L_0+1, L_0+1)-organization at the output $\tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)}$ between the minimum efficient scales. The average cost function of the (L_0, L_0+1)-organization crosses the average cost function of the (L_0+1, L_0+1)-organization at a higher level of output $\tilde{q}^{(L_0,L_0+1)\to(L_0+1,L_0+1)} > \tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)}$.

As a result, the ME firm with a (L_0, L_0)-organization adds a layer of middle managers at the headquarters or the establishment at the output $\tilde{q}^{*(L_0,L_0)}$ and a layer at the other unit at output $\tilde{q}^{(L_0,L_0+1)\to(L_0+1,L_0+1)} \in (\tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)}, \tilde{q}^{*(L_0+1,L_0+1)})$.

Proof. See Appendix C.3.6. □

Proposition 5 is a key result of the model. It states that the ME firm successively reorganizes the headquarters and the establishment as it grows. A firm with (L_0, L_0+1)-organization optimally combines two production technologies. At the output $\tilde{q}^{*(L_0,L_0)}$, the (L_0, L_0)-organization has the minimum average costs. The firm allocates total output and CEO time to the headquarters. For higher output $\tilde{q} > \tilde{q}^{*(L_0,L_0)}$, the average costs of the (L_0, L_0)-organization increase. The average costs of the (L_0, L_0+1)-organization decrease up to the minimum efficient scale of the $(L_0+1, L_0+$

1)-organization, because the firm allocates an increasing share of output to the establishment. Figure IVb illustrates this result.[9]

The managerial layer at the establishment releases CEO time: relative to output, the CEO spends a larger share of time at the headquarters than at the establishment. This keeps below-CEO knowledge low. The layer thus increases efficiency both at the establishment and the headquarters. It decreases the need to add a managerial layer at the headquarters. As in Propositions 3 and 4, the organization of a ME firm is interdependent: The optimal number of layers at the headquarters depends on the number of layers at the establishment (and vice versa).

Transport and helping cost frictions, $\tau > 1$, $\theta_{10} \geq \theta_{00}$, $w_1 = w_0$, $\tilde{q}_1 = \tilde{q}_0$. Figure Va illustrates the average production costs of different organizational structures if local output, wages and helping costs are equal across locations, but there are transport frictions. The average production costs are U-shaped, as those of a SE firm. This reflects how reallocating output is efficient only under certain conditions, and that higher output has a similar effect on firm organization as in a SE firm. Although the transport frictions affect the shape of the average cost function of the $(0,1)$-organization, they do not affect the pattern of reorganization. The ME firm first adds a layer at one unit and then the other.

Figure Vb illustrates how the helping costs across space θ_{10} affect the number of managerial layers of the firm. Higher helping costs increase the knowledge levels of employees and thus the marginal production costs at the establishment. Adding a layer helps the firm to mitigate the cost increase, because it allows decreasing production worker knowledge. The higher the helping costs, the smaller is the level of output at which the firm adds a layer at the establishment.

In addition to their effect on the number of layers, higher helping costs affect ME production per se. Higher helping costs reduce the desirability of maintaining an establishment relative to shipping output from the headquarters, as they increase the marginal production costs at the establishment.

Wage differences, $w_1 \neq w_0$, and output differences, $\tilde{q}_1 \neq \tilde{q}_0$. Appendix C.3.7 shows how wage and output differences between locations affect the optimal number and location of managerial

[9] The average cost function of the $(0,1)$-organization coincides with the average cost functions of the $(0,0)$-organization and the $(1,1)$-organization for quantities below and above the minimum efficient scales respectively, because for those levels of output, single establishment production with 0 and 1 below-CEO layers is more efficient than production with the $(0,1)$-organization.

layers. Most notably, higher wages or lower output at the establishment than at the headquarters can make it optimal to hire middle managers at the headquarters, but not the establishment. Output, wages and the helping costs across space jointly determine at which level of output it is optimal to hire middle managers at the headquarters, the establishment, or both units. The level decreases with higher helping costs.

IV.D. The optimal output

We return to the setting with many firms i outlined at the beginning of section IV.A. We assume that each firm faces a downward sloping demand curve for its product. Firms compete monopolistically, so there is no strategic interaction between firms. Firms choose output levels $\tilde{q}_j$ to maximize profits:

$$\max_{\tilde{q}_0, \tilde{q}_1 \geq 0} \pi_i = \sum_{j=0}^{1} p_j(\tilde{q}_j)\tilde{q}_j - C(\tilde{q}_0, \tilde{q}_1) \tag{25}$$

Proposition 6. *Suppose that the firm produces at the headquarters and the establishment. Suppose that the local production quantities are equal to local output (i.e.,* $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$*) and that they are sufficiently large. Higher helping costs across space* θ_{10} *decrease the optimal output at the establishment* $\tilde{q}_1$ *and increase the optimal output at the headquarters* $\tilde{q}_0$*.*

Proof. See Appendix C.4.[10] □

Higher helping costs across space decrease the establishment output and increase the headquarter output due to their effect on the marginal production costs. Consequently, the helping costs have both a direct effect on the managerial organization and an indirect effect through endogenous output. Thus, higher helping costs can increase the number of managerial layers both at the establishment and the headquarters due to the changes in output, as Figure C.5 illustrates.

IV.E. Comparison of facts and model

In the model, the helping costs θ_{10} reflect distance and other geographic frictions. Higher helping costs increase the marginal production costs of an establishment, and decrease its optimal size and

[10]The Appendix includes the results for the case that $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$ not considered in Proposition 6.

the attractiveness of a location for an establishment, consistent with Fact 1.

In line with Fact 2, the higher marginal costs at the establishment increase the use of middle managers there. Depending on local wages and local output, higher helping costs also increase the use of middle managers at the headquarters due to the common CEO.

Hiring middle managers at the establishment (or the headquarters) releases CEO time that is reallocated across locations and reduces the need for middle managers at the headquarters (or establishment). Firms thus successively add middle managers at the headquarters or the establishment as they grow, consistent with Fact 3.

V. Reorganization due to high-speed railway routes

V.A. Model predictions

We use the opening of high-speed railway routes (henceforth HSR) to provide evidence on the prediction that geographic frictions between an establishment and the headquarters affect not only the organization of the establishment, but also the organization of the headquarters and possible other establishments. Figure VI illustrates the model predictions using a directed graph. Solid circles denote observable variables, hollow circles denote unobservable variables, and the arrows denote causal links between variables. To keep the graph simple, we group firm and establishment-level variables and use semi-solid circles if only part of the group is observable.

The HSR exogenously reduce the travel times between an establishment k and the headquarters. In the terms of the model, lower travel times decrease the helping costs θ_{k0}. Lower helping costs increase establishment output $\tilde{q}_k$ and decrease headquarter output $\tilde{q}_0$. They thus have direct and indirect effects on firm organization. The effects often point in different directions. Lower helping costs directly affect the organizational structure ω by decreasing the optimal number of layers and increasing the attractiveness of maintaining an establishment. They indirectly affect the organizational structure because higher output increases the optimal number of layers.

Similarly, CEO knowledge $\bar{z}_{0,\omega}$, the allocation of CEO time $s_{j,\omega}$ and the allocation of production $q_{j,\omega}$ depend directly on θ_{k0}, but also indirectly through $\tilde{q}_j$ and ω. The number and knowledge of employees per layer $n^{\ell}_{j,\omega}, z^{\ell}_{j,\omega}$ depend directly on θ_{k0} and indirectly through $\bar{z}_{0,\omega}, s_{j,\omega}, q_{j,\omega}$ and ω. Given the organizational structure ω, lower helping costs unambiguously increase the number of production workers and the share of CEO time at the establishment, but the effect on most other variables is ambiguous (see Table D.1). Endogenous changes of ω increase ambiguity, because knowledge changes discontinuously if firms change layers (as in CRH, for a proof, see Appendix C.3.6).

The complexity of the relationship between the helping costs θ_{k0} and firm organization has implications for the interpretation of the empirical estimates. The results in Proposition 4 hold conditional on output and the organizational structure. These variables do not vary exogenously, but depend on the helping costs, and we do not have instruments for them. If we conditioned on output or organizational structure in an establishment-level regression, the estimation would entail a "bad control" problem (Angrist and Pischke 2008, 64-68). Our empirical exercise therefore estimates the total—direct and indirect—effect of changes in helping costs using reduced-form regression equations.

V.B. Travel time data

We use data on the travel times between the 115 German train stations connected to the long-distance railway network from Deutsche Bahn AG, the state-owned railway firm. Travel times changed substantially due to the opening of three HSR during our sample period.[11] Figure VII displays a map of the new HSR and how they connect to the existing long-distance network. Trains on these routes exclusively transport people. Except for route 2, the high-speed trains run at up to 300 km/h, which is around 100 km/h faster than on the other routes of the German long-distance network. Appendix D.2 provides details on the routes and their construction.

As Figure VII shows, the German railway network is highly interconnected compared to that of other countries. For instance, the French railway network approximately has a "star" structure with Paris as the center. The German network features several hubs. The HSR therefore affect more cities than merely those at the immediate ends. For example, route 1 between Cologne and Frankfurt reduced travel times from cities in the Ruhr area to those in East and South Germany, such as Leipzig, Stuttgart, and Würzburg.

We use data on the minimum net travel times and the number of changes between cities in the years 2000, 2004 and 2008. We follow Deutsche Bahn AG and compute travel times as time on

[11] A fourth route between Leipzig and Berlin opened in 2006. However, the travel time between these cities decreased only gradually according to the data, so the route is not used in the estimation.

the train plus 30 minutes per change. We merge the travel times and the firm data based on the county where the establishment and the station is located. We restrict the sample to firms that have headquarters and at least one establishment connected to the long-distance network to avoid unobservable differences between connected and unconnected firms driving the results. Table D.3 displays summary statistics for our sample.

A possible concern is that trains are not an attractive means of transportation for business travelers. However, this is not true of the high-speed trains. According to information from Deutsche Bahn AG for the year 2017, the share of business travelers on the new routes was about double their average share.[12] This is unsurprising given that the HSR render the train the fastest means of transportation between the connected cities. It is faster to travel by train than by car—it takes almost twice as long to drive from Frankfurt to Cologne, for example—or even plane. In addition, the high-speed trains are a flexible means of travel as regular tickets are valid on all trains that service a connection.

V.C. Empirical specification

To gauge the effect of lower travel times on directly affected establishments, we estimate:[13]

$$y_{ijt} = \beta_0 + \beta_1 \mathbf{1}\{\textit{Lower travel times to HQ}\}_{ijt} + \alpha_j + \alpha_{ct} + \epsilon_{ijt} \tag{26}$$

i refers to a ME firm, j to an establishment, c to the county where an establishment is located and t indexes time. α denotes fixed effects. The variable of interest is an indicator variable for a travel time reduction between the establishment and its headquarters of at least 30 minutes.

To understand the effect on the headquarters, we estimate:

$$y_{iht} = \beta_0 + \beta_1 \mathbf{1}\{\exists j \textit{ with lower travel times to HQ}\}_{iht} + \alpha_h + \alpha_{dt} + \epsilon_{iht} \tag{27}$$

h denotes the headquarters and d the headquarter county. The variable of interest indicates if travel times to at least one establishment decrease by at least 30 minutes.

[12] The statistics are computed based on the fraction of tickets sold with a corporate discount.
[13] This specification is similar to Charnoz et al. (2018).

To assess the effect on non-directly affected establishments of affected firms, we estimate:

$$\begin{aligned}(28)\qquad y_{ikt} = \beta_0 + \beta_1 \mathbf{1}\{&\textit{No lower travel times to HQ}\\ &\wedge \exists j \neq k \textit{ with lower travel times to HQ}\}_{ikt} + \alpha_k + \alpha_{ct} + \alpha_{dt} + \epsilon_{ikt}, \quad k \neq j\end{aligned}$$

k refers to a non-directly affected establishment. The indicator variable is equal to one if the travel time between establishment k and the headquarters does not decrease by at least 30 minutes, but the travel time between one of the other establishments of the firm and the headquarters does. As outcome variables $y_{i.t}$, we use the number of non-managerial employees to measure size, the wages of non-managerial employees to approximate knowledge, and the number of managerial layers, complemented with the managerial share, to measure organization.

We set the indicators equal to one if the travel time between an establishment and the headquarters decreases by at least 30 minutes because the HSR decrease the travel times by at least 30 minutes. The threshold thus helps us to ensure that the reduction is driven by the exogenous new HSR instead of potentially endogenous demand-driven adjustments to the time-table.[14]

The specifications mimic difference-in-differences estimation. The "treatment" is lower travel times between the directly affected establishment and the headquarters (eq. 26), or between at least one establishment and the headquarters (eq. 27, 28). Its baseline effect is captured by the establishment or headquarter fixed effects. The (headquarter) county $\times$ year fixed effects capture the "after" dummy. The indicator variables $\mathbf{1}\{\cdot\}$ correspond to the interaction of the "treatment" and "after" dummies. We implement the estimation using the `reghdfe` command by Correia (2014).

Lower travel times may affect other model parameters, such as local wages because employees commute longer distances (Heuermann and Schmieder 2019). Firms may also benefit from better suppliers (Bernard et al. 2019). The (headquarter) county $\times$ year fixed effects isolate the impact of lower geographic frictions on firm organization from other forces. Specifically, the regressions for directly affected establishments compare establishments with travel time reductions and establishments *in the same county and year* without reductions. Lower local wages or better suppliers

[14] One may worry that a possibly endogenous reduction in the number of changes triggers the treatment dummy. In the data, the number of changes decreases either due to the new HSR, or if a station is connected to the long-distance network. Our results are robust to restricting the sample to stations connected to the long-distance network in all years (see Table D.21).

benefit all establishments, so our estimation strategy accounts for their effect. Similarly, the regressions for the headquarters compare headquarters with travel time reductions to at least one establishment to headquarters *in the same county and year* without reductions. The specification for non-directly affected establishments compares establishments that belong to firms with treated establishments to establishments *in the same county and year* that belong to firms without treated establishments, additionally accounting for shocks at the headquarter location.[15] Being treated in this set-up presupposes that firms have at least two establishments, so we restrict the sample accordingly. Due to the fixed effects, only establishments and headquarters in counties with at least one affected unit identify the coefficient β_1. We hence drop counties without affected units.

If the effects of possible omitted variables vary non-linearly with unit size, the (headquarter) county $\times$ year fixed effects may not fully absorb them. We hence match treated and control units by size before treatment, using the average number of employees in 2000/2001 as size measure. We employ the Coarsened Exact Matching algorithm (Iacus et al. 2012), because its weighting procedure helps estimate the average treatment effect on the treated. In addition to the size quartile, we match units by year and (headquarter) county to mimic the fixed effects in the regressions.[16]

We address several possible concerns with respect to our identification strategy. Importantly, we hypothesize based on the model that lower travel times have an effect on non-directly affected establishments of affected firms. If these establishments are located in the same counties as directly affected establishments, they may contaminate the control group in equation (26). We re-run regressions for firms with only one establishment and, alternatively, exclude non-directly affected establishments of affected firms from the sample to account for this possibility.

A second important concern is that firms may strategically locate their establishments close to the HSR in anticipation of their opening. To address this possibility, we re-run regressions for establishments set up before 2000, the first year of the sample, and before 1995, when construction of route 1 started. A few establishments and headquarters move from one county to another during

[15] The strictest specification would condition on county $\times$ headquarter county $\times$ year fixed effects, i.e., compare non-directly affected and unaffected establishments in the same county with headquarters in the same county. However, there are too few such pairs in the sample to run these regressions.

[16] In a few robustness checks, matching by county makes the size of treated and control units *less* similar due to the uneven spatial distribution of units. In these cases, we match units only by size quartile and year. We would ideally like to match non-directly affected establishments by headquarter county, county and year, but there are too few pairs in the sample to do so. We match establishments by size quartile and year in Table VII and report results for matching on size quartile, (headquarter) county, and year in Appendix Table D.12.

the sample period. We use their original location for the main analyses and drop them from the sample in robustness checks (Table D.20).

Finally, we document that the outcomes of treated and control units follow similar trends before the opening of the routes. This supports our assumption that the two groups differ only with respect to the travel time changes.

V.D. Regression results

Table VII presents the regression results for the 2000-2010 panel. Columns 1 to 4 contain results for all firms. Columns 5 to 8 restrict the sample to firms with at least two establishments.

As the top panel shows, lower travel times increase the size of the directly affected establishments. The number of non-managerial employees increases by eight percent. This increase in size is not accompanied by an increase in wages, the number of layers or the managerial share. The middle panel shows that lower travel times lead to organizational adjustments at the headquarters of firms with at least two establishments. Average wages of non-managerial employees increase by two percent. The number of managerial layers and the managerial share also increase significantly. The coefficient estimate is equivalent to an increase of the managerial share by nine percent in the average firm. As the bottom panel shows, the impact of lower travel times goes beyond the headquarters and the directly affected establishment. Both wages and the managerial share increase at establishments that do not themselves benefit from lower travel times, but belong to firms that do.

Overall, the results strongly support the prediction of the model that geographic frictions between an establishment and the headquarters affect the organization of not only the establishment, but also the headquarters and possible other establishments of the firm. The results are consistent with the interpretation that lower helping costs due to faster travel times improve the establishment's access to the CEO or, more generally, managerial resources of the headquarters. This allows the establishment to grow without local organizational adjustments. Instead, the firm increases managerial capacity at the headquarters. Through the lens of the model, both the higher non-managerial wages, reflecting higher knowledge, and the higher number of layers at the headquarters reflect adjustments to release CEO time. The adjustments at the non-directly affected establishments support the interpretation that the firm reallocates managerial resources of the headquarters from non-directly affected to the directly affected establishments.

Evidence supporting the validity of the identification strategy. Figure VIII shows that the effect of lower travel times on the size of directly affected establishments is similar when we exclude possibly non-directly affected establishments from the sample and account for possible strategic location of establishments. The coefficients vary, but are not significantly different from the baseline effect. The larger coefficient for firms with one establishment possibly reflects that these establishments are smaller than those in the baseline sample. The smaller coefficient for the sample without indirectly affected establishments may reflect that those grow more slowly than establishments of unaffected firms. Table D.4 displays the complete set of regression results.

Figure IX documents that the outcomes of treated and control units follow similar trends before the opening of the HSR. This supports the assumption that the control units provide a valid counterfactual for the treated units after treatment. Table D.5 contains the results for all outcomes and samples as well as the indirectly affected establishments.

Robustness. Appendix D.5 documents that the results are robust to alternative approaches to statistical inference, alternative variable definitions, and alternative sample restrictions. Notably, we find that the education and experience of employees change concomitantly with wages, in line with our knowledge hierarchy model. The model proposes that lower travel times affect firm organization via the specific channel of lower helping costs. To support this channel, Appendix D.6 documents that the effects are stronger in sectors with a less predictable production process.

VI. Conclusion

This paper showed that the managerial organization of ME firms is interdependent across establishments. Specifically, we showed empirically and theoretically that geographic frictions between an establishment and the headquarters not only affect the organization of this particular establishment, but also the organization of the headquarters and other establishments of the firm.

Our paper opens up several avenues for future research. In one direction, future work could study the productivity effects of reorganization in ME firms (as done for SE firms by Caliendo et al. 2020). Quantifying the local and non-local productivity effects of establishment reorganization would improve our understanding of firm performance. It would also lay the foundation for work on the propagation of shocks across space through firm organization and a comparison of the importance of managerial and financial constraints as propagation mechanisms.

In another direction, future work could exploit ME firms as a setting that opens up new angles on within-firm processes and frictions. For example, it has been difficult to empirically differentiate the knowledge and the monitoring hierarchy models, because they yield very similar predictions for SE firm organization (Chen and Suen 2019). The ME firm setting may provide an opportunity to shed light on the nuances between the two frameworks. Our own preliminary analyses suggest that a monitoring hierarchy model, where geographic frictions make it more difficult to monitor workers at the establishment, yields predictions that are similar to those of our model for establishment organization, but that differ with respect to headquarter organization. Testing these differences is unfortunately beyond the scope of our data, but would be very valuable for our understanding of firm organization, both with one and multiple establishments.

LMU Munich, CEPR and CESifo,

LMU Munich and Stanford Graduate School of Business, and

Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB).

References

Angrist, Joshua D. and Jörn-Steffen Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Antràs, Pol, Esteban Rossi-Hansberg, and Luis Garicano (2008). Organizing Offshoring: Middle Managers and Communication Costs. In E. Helpman, T. Verdier, and D. Marin (Eds.), *The Organization of Firms in a Global Economy*, pp. 311–339. Cambridge, MA: Harvard University Press.

Antràs, Pol and Stephen R Yeaple (2014). Multinational Firms and the Structure of International Trade. *Handbook of International Economics 4*, 55–130.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson (2014). Vertical integration and input flows. *American Economic Review 104*(4), 1120–48.

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier (2020). Face-to-Face Communication in Organisations. *Review of Economic Studies, forthcoming*.

Bernard, Andrew B, Andreas Moxnes, and Yukiko U Saito (2019). Production networks, geography and firm performance. *Journal of Political Economy 127*(2), 639–688.

Bertrand, Marianne (2009). CEOs. *Annual Review of Economics 1*(1), 121–150.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Italy Saporta-Eksten, and John Van Reenen (2019). What drives differences in management practices? *American Economic Review 109*(5), 1648–83.

Blossfeld, Hans-Peter (1983). Höherqualifizierung und Verdrängung-Konsequenzen der Bildungsexpansion in den siebziger Jahren. In M. Haller and W. Müller (Eds.), *Beschäftigungssystem im gesellschaftlichen Wandel*, pp. 184–241. Frankfurt/New-York: Campus-Verlag.

Blossfeld, Hans-Peter (1987). Labor-market entry and the sexual segregation of careers in the Federal Republic of Germany. *American Journal of Sociology 93*(1), 89–118.

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg (2020). Productivity and organization in Portuguese firms. *Journal of Political Economy 128*(11), 4211–4257.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg (2015). The Anatomy of French Production Hierarchies. *The Journal of Political Economy 123*(4), 809–852.

Caliendo, Lorenzo and Esteban Rossi-Hansberg (2012). The impact of trade on organization and productivity. *The Quarterly Journal of Economics 127*(3), 1393–1467.

Card, David, Jörg Heining, and Patrick Kline (2013). Workplace heterogeneity and the rise of West German inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Charnoz, Pauline, Claire Lelarge, and Corentin Trevien (2018). Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail. *The Economic Journal 128*(610), 949–994.

Chen, Cheng (2017). Management Quality and Firm Hierarchy in Industry Equilibrium. *American Economic Journal: Microeconomics 9*(4), 203–244.

Chen, Cheng and Wing Suen (2019). The Comparative Statics of Optimal Hierarchies. *American Economic Journal: Microeconomics 11*(2), 1–25.

Correia, Sergio (2014). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Statistical Software Components S457874, Boston College Department of Economics*.

Crèmer, Jacques, Luis Garicano, and Andrea Prat (2007). Language and the theory of the firm. *The Quarterly Journal of Economics 122*(1), 373–407.

Friedrich, Benjamin U. (2020). Trade Shocks, Firm Hierarchies, and Wage Inequality. *Review of Economics and Statistics*, forthcoming.

Garicano, Luis (2000). Hierarchies and the organization of knowledge in production. *The Journal of Political Economy 108*(5), 874–904.

Garicano, Luis and Esteban Rossi-Hansberg (2015). Knowledge-based hierarchies: Using organizations to understand the economy. *Annual Review of Economics 7*(1), 1–30.

Giroud, Xavier (2013). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics 128*(2), 861–915.

Giroud, Xavier and Holger M. Mueller (2019). Firms' Internal Networks and Local Economic Shocks. *American Economic Review 109*(10), 3617–49.

Guadalupe, Maria and Julie Wulf (2010). The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies. *American Economic Journal: Applied Economics 2*(4), 105–127.

Gumpert, Anna (2018). The organization of knowledge in multinational firms. *Journal of the European Economic Association 16*(6), 1929–1976.

Helpman, Elhanan, Marc J. Melitz, and Stephen Ross Yeaple (2004). Export versus FDI with heterogeneous firms. *The American Economic Review 94*(1), 300–316.

Heuermann, Daniel F. and Johannes F. Schmieder (2019). The effect of infrastructure on worker mobility: evidence from high-speed rail expansion in Germany. *Journal of Economic Geography 19*(2), 335–372.

Iacus, Stefano M., Gary King, and Giuseppe Porro (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis 20*(1), 1–24.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla (2013). The margins of multinational production and the role of intrafirm trade. *Journal of Political Economy 121*(1), 74–126.

Kalnins, Arturs and Francine Lafontaine (2013). Too far away? the effect of distance to headquarters on business establishment performance. *American Economic Journal: Microeconomics 5*(3), 157–179.

Keller, Wolfgang and Stephen R. Yeaple (2013). The Gravity of Knowledge. *The American Economic Review 103*, 1414–1444.

Mariscal, Asier (2018). Firm Organization and Information Technology: Micro and Macro Implications. *mimeo, Universidad Carlos III-Madrid*.

McElheran, Kristina (2014). Delegation in Multi-Establishment Firms: Evidence from IT Purchasing. *Journal of Economics & Management Strategy 23*(2), 225–258.

Papke, Leslie E. and Jeffrey M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Rajan, Raghuram G. and Julie Wulf (2006). The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies. *The Review of Economics and Statistics 88*(4), 759–773.

Seetharam, Ishuwar (2018). The Indirect Effects of Hurricanes: Evidence from Firm Internal Networks. *mimeo, Stanford University*.

Sforza, Alessandro (2020). Shocks and the organization of the firm: Who pays the bill? *CESifo Working Paper No. 8084*.

Spanos, Grigorios (2019). Firm Organization and Productivity Across Locations. *Journal of Urban Economics 112*, 152–168.

Table I: Descriptive statistics, SE vs. ME firms, 2012 cross section

Units of observation	N	of which ME firms (% share)
Firms	109,357	9.0
with non-missing sales	57,811	9.1
Establishments	144,437	31.1
Employees	6,356,072	34.2

Descriptive statistics	N	ME	Mean	SD	p25	p50	p75	p95
# employees	99,545	0	42	92	13	21	39	133
	9,812	1	222	1,980	22	50	127	650
Sales (M €)	52,524	0	28	694	2	4	9	67
	5,287	1	358	4,111	4	15	74	608
# managerial layers	99,545	0	1.4	1.0	1	1	2	3
	9,812	1	2.0	1.0	1	2	3	3
Managerial share	99,545	0	28	28	5	19	43	88
(%, layers)	9,812	1	36	29	11	29	58	90
Managerial share	99,545	0	6	11	0	0	9	27
(%, Blossfeld)	9,812	1	9	12	0	5	12	33

Descriptive statistics. *ME:* indicator for multi-establishment firm; *# employees:* number of full-time employees; *Sales (M €)*: sales in million €; *# managerial layers*: number of managerial layers; *Managerial share (%, layers/Blossfeld):* share of wage sum earned by employees in managerial occupations (according to layers/Blossfeld occupational categories).

Table II: Descriptive statistics, ME firms, 2012 cross section

Descriptive statistics, firm	N		Mean	SD	p50	p75	p95
# establishments (incl. HQ)	9,812		4.6	19.6	2	3	10
Maximum distance to HQ, km	9,812		218	189	167	376	547
Minimum area covered, km^2	3,579		30,117	41,725	7,025	49,915	125,253
Descriptive statistics, HQ/ est.	**N**	**HQ**	**Mean**	**SD**	**p25**	**p50**	**p75**
# employees	35,080	0	32	333	2	5	16
	9,812	1	107	669	11	27	76
# managerial layers	35,080	0	1.0	0.8	0	1	2
	9,812	1	1.7	1.1	1	2	3
Managerial share	35,080	0	37	38	0	24	70
(%, layers)	9,812	1	38	32	11	32	64
Managerial share	35,080	0	8	19	0	0	5
(%, Blossfeld)	9,812	1	10	16	0	4	14

Descriptive statistics, ME firms, headquarters (HQ) and establishments (est.s). *# establishments (incl. HQ):* number of establishments (including headquarters); *Maximum distance to HQ, km:* maximum distance between establishments and headquarters in kilometers, where distance is computed as the population weighted average of the distances between all municipalities in the establishment county and the headquarter county; *Minimum area covered, km^2:* minimum area covered by establishments and headquarters in square-kilometers, only available for firms with at least two establishments in addition to HQ; *HQ:* indicator for headquarters; others see Table I.

Table III: Location probability and establishment size, ME firms, 2012 cross section

Dependent variable	Location probability			Log # est. employees		
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	-0.315^{***}	-0.303^{***}	-0.368^{***}	-0.106^{***}	-0.112^{***}	-0.137^{***}
	(0.021)	(0.023)	(0.020)	(0.018)	(0.019)	(0.017)
Log market potential	0.745^{***}	0.780^{***}		0.485^{***}	0.465^{***}	
	(0.026)	(0.031)		(0.044)	(0.046)	
Relative wages	-0.942^{***}	-0.887^{***}		-0.330^{**}	-0.433^{***}	
	(0.062)	(0.063)		(0.108)	(0.109)	
Relative land prices		-0.021^{***}			0.020^{***}	
		(0.005)			(0.005)	
# observations	3,715,666	3,222,108	3,715,666	21,496	19,203	21,496
# firms	9,266	8,732	9,266	3,006	2,773	3,006
HQ sector FE	Y	Y	Y	N	N	N
HQ county FE	Y	Y	Y	N	N	N
Legal form FE	Y	Y	Y	N	N	N
County FE	N	N	Y	N	N	Y
Firm FE	N	N	N	Y	Y	Y
Model		Probit			OLS	

The table presents the coefficient estimates of a Probit model in columns 1-3 (constant included; standard errors clustered by HQ county in parentheses) and a linear model in columns 4-6 (standard errors clustered by firm and county in parentheses). The regressions in columns 4 to 6 control for firm fixed effects, hence they only include ME firms with establishments in at least two counties. ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1)-(3): indicator for whether firm i owns at least one establishment in county c, (4)-(6): log number of employees of establishment(s) in county c. *Independent variables: Log distance to HQ:* log distance between county c and HQ county of firm i in km; *Log market potential:* log of average of GDP of county c and surrounding counties weighted by distance; *Relative wages/land prices:* average wages/land prices in county c relative to wages/land prices in HQ county of firm i. We compute average wages in a county excluding firm i. Distance, market potential and relative land prices are computed using data of the German Federal Statistical Office. The number of firms is lower than the number of ME firms due to missing values for the legal form. FE=fixed effects.

Table IV: Regression results, managerial organization of ME firms, 2012 cross section

Dependent variable	# managerial layers				Mg. share $\in [0,1]$ Layers		Mg. share $\in [0,1]$ Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.039^{***}		0.050^{***}		0.152^{***}		0.131^{***}	
	(0.007)		(0.006)		(0.014)		(0.019)	
Log area		0.025^{***}		0.027^{***}		0.070^{***}		0.072^{***}
		(0.005)		(0.004)		(0.011)		(0.013)
Log sales	0.115^{***}	0.082^{***}						
	(0.004)	(0.006)						
Log # non-mg. employees			0.109^{***}	0.088^{***}				
			(0.004)	(0.005)				
# firms	4,323	1,661	7,742	2,768	7,742	2,768	7,742	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. *** p < 0.001. Even columns include only ME firms with establishments in at least two counties. The generalized linear model (GLM) assumes a logit link function and the binomial distributional family. *Dependent variable:* (1)-(4) number of managerial layers, (5),(6) managerial share in wage sum, according to layers, (7),(8) managerial share in wage sum, according to Blossfeld occupational categories. *Independent variables: Maximum log distance to headquarters*: log of maximum distance between establishment and headquarters in km; *Log area spanned by firm*: log of minimum area covered by establishments and headquarters in square kilometers; *Log sales:* log annual sales; *Log # of non-mg. employees:* log number of employees at lowest layer. FE = fixed effects, mg. = managerial.

Table V: Regression results, mg. organization of establishments, 2012 cross section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0, 1]$		# layers	Mg. share $\in [0, 1]$	
		Layers			Layers	
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance	0.056***	0.191***	0.236***			
to HQ	(0.011)	(0.032)	(0.053)			
Maximum log				0.057***	0.166***	0.157***
distance to HQ				(0.007)	(0.016)	(0.022)
Log # non-mg.	0.256***			0.162***		
employees	(0.010)			(0.004)		
# est./HQ	26,409	31,717	31,717	7,999	8,217	8,217
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). ** $p < 0.01$, *** $p < 0.001$. The generalized linear model (GLM) assumes a logit link function and the binomial distributional family. *Dependent variable:* (1),(4) number of managerial layers, (2),(5) managerial share in wage sum, according to layers, (3),(6) managerial share in wage sum, according to Blossfeld occupational categories. *Independent variables: Log distance to headquarters*: log of distance between establishment and headquarters in km; *Maximum log distance to headquarters*: log of maximum distance between establishment and headquarters in km; *Log # of non-mg. employees:* log number of employees at lowest layer in establishment/HQ. FE = fixed effects, mg. = managerial.

Table VI: Transition dynamics of the managerial organization, 2000–2010 data

(a) # managerial layers of firm

# layers in $t/t+1$	0	1	2	3	SE	# firms
0	**85**	8	1		6	10,778
1	5	**81**	8	1	6	18,274
2		7	**79**	8	5	18,754
3			6	**90**	4	22,391

(b) # managerial layers at headquarters/establishment(s)

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**85**	5						6	10,778
HQ 1/ est. 0	6	**75**	4	6				8	8,340
HQ 1/ est. 1	1	5	**76**	7		1		4	8,052
HQ 2/ est. 0,1		4	4	**76**	2	6		7	12,046
HQ 2/ est. 2			1	10	**69**	9	1	2	3,410
HQ 3/ est. 0,1,2				5	2	**84**	3	5	13,365
HQ 3/ est. 3						9	**86**	1	4,625

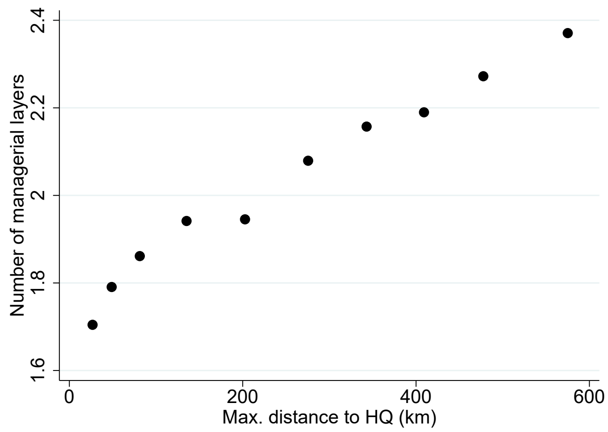
Panel (a) displays the percentage share of firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers or to SE firm status in year $t+1$ (given in the columns). Panel (b) displays the percentage share of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization or to SE firm status in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the HQ. The figure behind the slash denotes the maximum number of layers of the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. Empty cells contain fewer than .5% of firms. Fewer than .5% of firms exit. Diagonal in bold.

Table VII: Lower travel times affect all units of ME firms, 2000–2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.084***	0.002	0.009	0.059	0.083***	0.004	0.016	0.049
	(0.019)	(0.004)	(0.016)	(0.253)	(0.021)	(0.005)	(0.018)	(0.273)
# observations	47,732	47,732	47,732	47,732	40,143	40,143	40,143	40,143
# est.	5,609	5,609	5,609	5,609	4,791	4,791	4,791	4,791
R-squared	0.901	0.931	0.875	0.890	0.899	0.932	0.878	0.902
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.013	0.004	0.019	0.103	0.023	0.018*	0.065**	0.994+
	(0.019)	(0.005)	(0.022)	(0.338)	(0.033)	(0.008)	(0.022)	(0.504)
# observations	13,393	13,393	13,393	13,393	6,261	6,261	6,261	6,261
# HQ	1,469	1,469	1,469	1,469	683	683	683	683
R-squared	0.951	0.951	0.875	0.913	0.951	0.945	0.872	0.919
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.017	0.019***	0.020	0.521*
					(0.020)	(0.004)	(0.017)	(0.209)
# observations					45,508	45,508	45,508	45,508
# est.					5,508	5,508	5,508	5,508
R-squared					0.917	0.926	0.890	0.887
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (headquarter) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *wages:* average log wages of non-managerial employees; *# lay.:* number of managerial layers; *mg.sh.:* share of managerial occupations in wage sum in percent, where managerial occupations are determined according to Blossfeld occupational categories. Treated and control units are matched by size quartile, (headquarter) county and year, except for the non-directly affected establishments that are matched by size quartile and year. All variables are winsorized at the first and 99th percentiles. The P-value for the managerial share of the headquarters in column 8 is 5.4%.

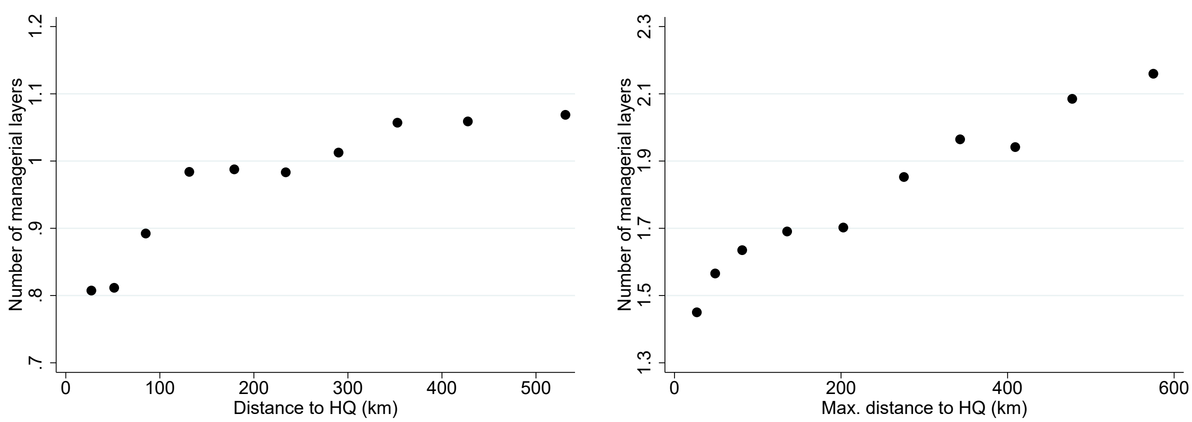
Figure I: Distance to HQ correlates positively with number of managerial layers of ME firms



Bin scatter plot of the relation between the maximum distance of establishments to headquarters (HQ) and the number of managerial layers in a multi-establishment (ME) firm. 2012 cross section, # firms: 8,217. Firms with establishments only in the headquarter county are excluded for consistency with Table IV.

Figure II: Distance to HQ correlates positively with number of managerial layers of HQ and establishments

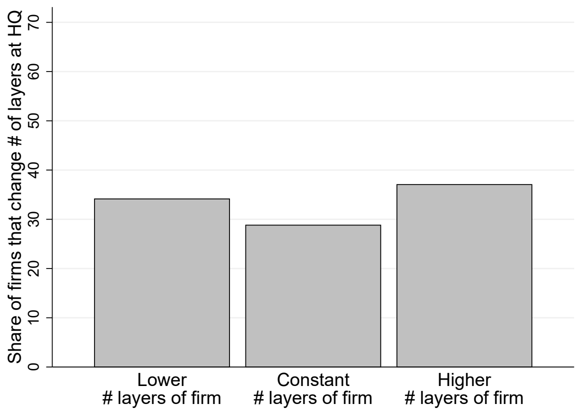
(a) Establishments (b) Headquarters



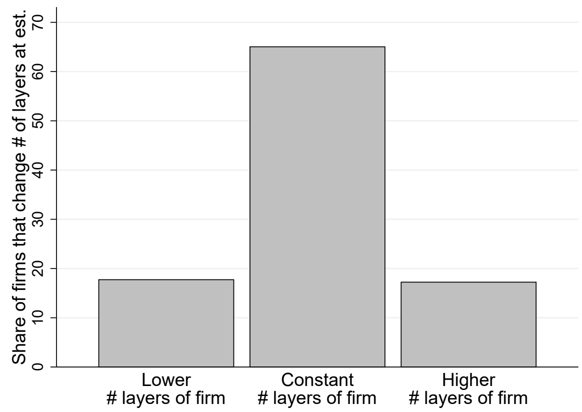
Bin scatter plot of the relation between (a) the distance of an establishment to the headquarters and the number of managerial layers of the establishment, and (b) the maximum distance of the establishment(s) to the headquarters and the number of managerial layers of the headquarters. 2012 cross section, # establishments: 31,718, # headquarters: 8,217. Establishments located in the headquarter county and headquarters of ME firms with establishments only in the headquarter county are excluded for consistency with Table V.

Figure III: Reorganization at the unit level often not visible at the firm level

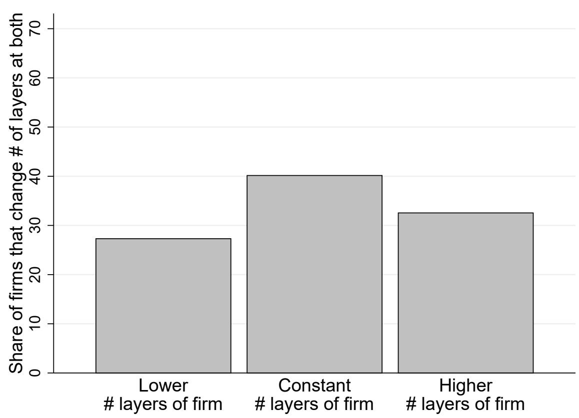
(a) # of layers changes only at headquarters



(b) # of layers changes only at establishments



(c) # of layers changes at headquarters and establishments

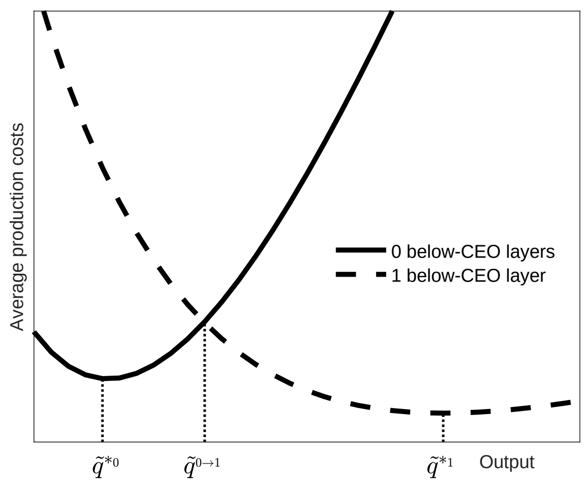


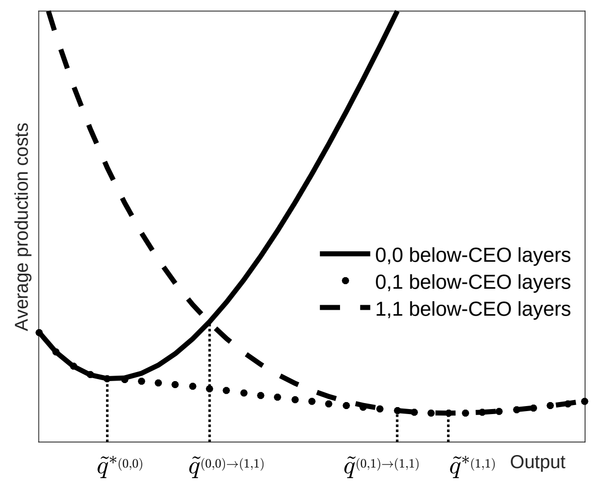
The figure shows that reorganization at the unit level is often not visible as reorganization at the firm level. The sample contains 13,177 firms that change the number of layers at the headquarters or the establishment(s) in the 2000–2010 panel. 49% (42%) of firms change the number of layers only at the headquarters (establishments). 9% of firms change the number of layers at both units.

Figure IV: Illustration of the average cost function, no transport frictions

(a) SE firm

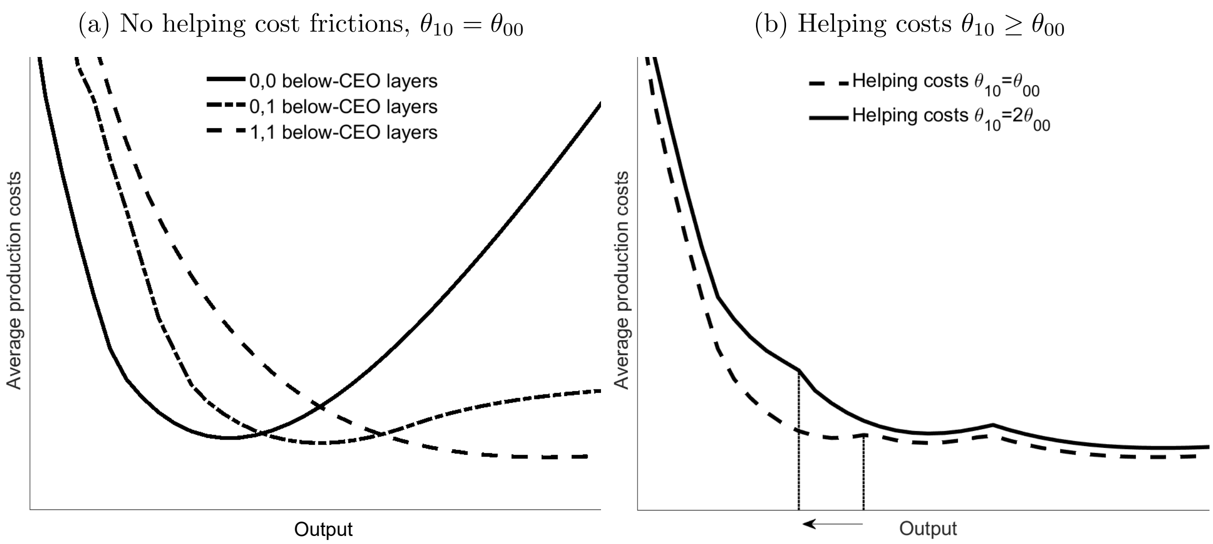
(b) ME firm





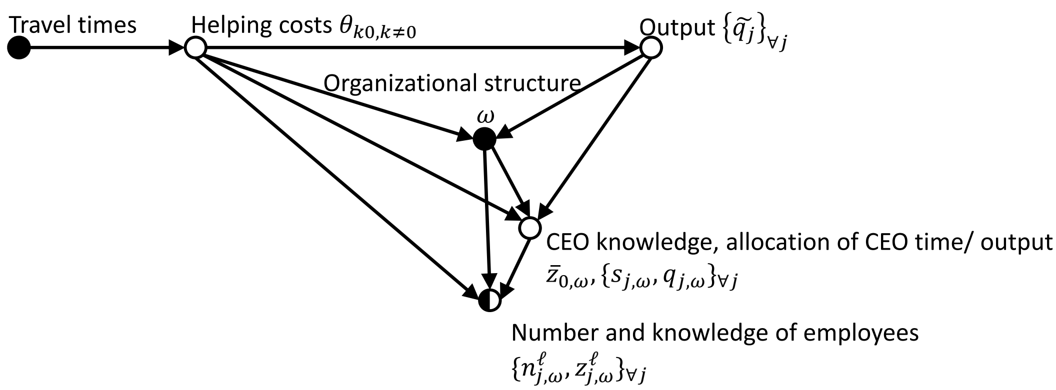
The figure plots the average cost functions of a SE and a ME firm for $\tau = 1$, $w_0 = w_1$, $\theta_{00} = \theta_{10}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from CRH), $w_0 = 1$. (a): The average cost function of a SE firm is U-shaped given L. The firm adds a layer at $\tilde{q}^{0\to1}$. (b): The average cost function of a ME firm with a symmetric number of below-CEO layers is U-shaped. The firm adds a layer at the establishment at $\tilde{q}^{*(0,0)}$ and at the headquarters at $\tilde{q}^{(0,1)\to(1,1)} > \tilde{q}^{(0,0)\to(1,1)}$ (or vice versa, as the $(0,1)$ and the $(1,0)$-organization have the same costs).

Figure V: Illustration of the average cost functions, transport frictions



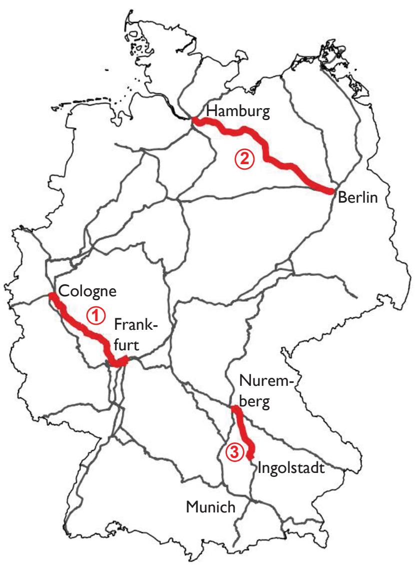
The figure plots average cost functions of a ME firm. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from CRH), $w_0 = 1$, $w_1 = w_0$, $\tau = 1.1$, $\tilde{q}_0 = \tilde{q}_1$. (a): At each kink, the ME firm adds a layer at one unit. The $(0,1)$ and $(1,0)$-organization have the same costs. (b): Higher helping costs θ_{10} decrease the output at which the firm reorganizes.

Figure VI: Model predictions regarding the effects of a change in the travel times



The graph illustrates the predictions of the model regarding the effects of a change in travel times. The arrows denote causal relationships between the variables at the nodes. The node symbol $\bullet$ ($\circ$) denotes that a variable is (un)observable. ◐ denotes that a group of variables contains observable and unobservable variables.

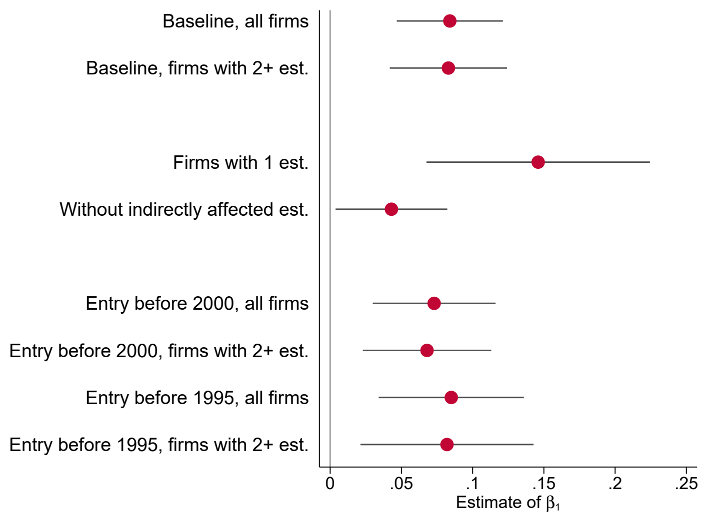
Figure VII: The new high-speed railway routes and the German long-distance network



Route	Date of opening	Travel time before	Travel time after	Change
1) Cologne—Frankfurt	Aug. 2002	135 min	76 min	-44%
2) Berlin—Hamburg	Dec. 2004	135 min	90 min	-33%
3) Ingolstadt—Nuremberg	May 2006	66 min	30 min	-55%

The map shows the German long-distance rail network (black) and the new high-speed railway routes (bold red). Data from Deutsche Bahn AG (http://data.deutschebahn.com/dataset/geo-strecke).

Figure VIII: Robustness to alternative control groups and strategic location of establishments

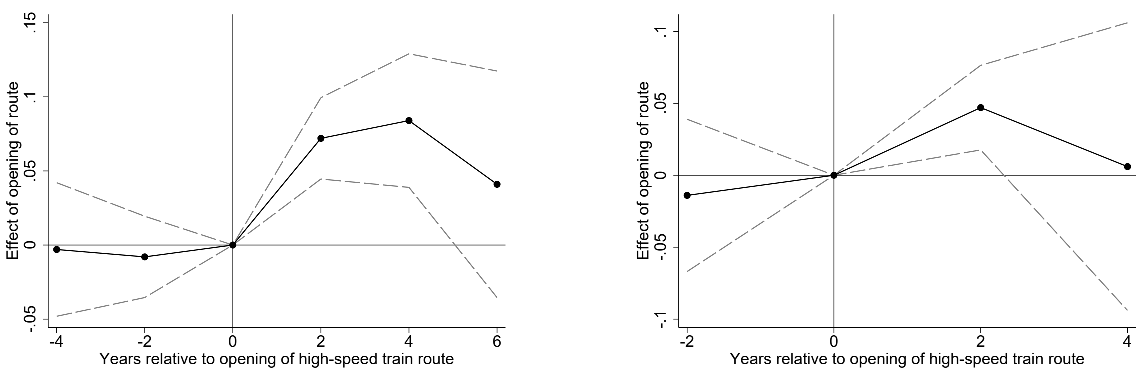


The figure plots the coefficients and 95% confidence intervals for the size of directly affected establishments from Table VII as well as from regressions on the sample of firms with one establishment, the sample excluding non-directly affected establishments, and samples varying the entry year. Treated and control establishments are matched on size quartile and year for the samples of firms with one establishment and without non-directly affected establishments, and on size quartile, county and year in the other cases. Standard errors clustered by county. Table D.4 displays the regression results.

Figure IX: The effect of the opening of high-speed railway routes

(a) Log # non-mg. employees, directly affected establishment

(b) # managerial layers, headquarters



The figure plots coefficients and 95% confidence intervals for regressions similar to equations (26) and (27). The dependent variable is (a) the log number of non-managerial employees of an establishment and (b) the number of managerial layers of HQ. The explanatory variables are indicator variables for (a) lower travel times to the HQ and (b) lower travel times to at least one establishment, interacted with biannual fixed effects. The excluded interaction is the year of the opening of the HSR. We control for establishment (HQ) fixed effects and (HQ) county $\times$ year fixed effects. We consider a shorter time period for HQ due to a low number of observations. Sample restricted to firms with at least two establishments. Standard errors are clustered by (HQ) county. Table D.5 displays the regression results.

FIRM ORGANIZATION WITH MULTIPLE ESTABLISHMENTS

ONLINE APPENDIX

Anna Gumpert

Henrike Steimer

Manfred Antoni

Table of Contents

A. Data

A.1. Data sources and record linkage procedure

A.1.1. Social security data

Employee history. The Integrated Employment Biographies (*Integrierte Erwerbsbiografien, IEB*) are based on records from the German Social Security System. They contain information on all employees subject to social insurance contributions since 1975 and are updated at least annually. The data cover nearly all private sector employees in Germany, but do not include civil servants and the self-employed. The IEB contain information on birth year, gender, nationality, education, occupation, full time or part-time status and daily earnings of each employee. Daily earnings are censored at the social security limit. Jacobebbinghaus and Seth (2007) and Antoni et al. (2016) provide a detailed description of the structure of the data.[17]

Information on education is not reported in all periods for every individual, but can be inferred from other observations on the same individual. We follow Fitzenberger et al. (2005) and impute missing values of the education variable based on past and future information.

Establishment History Panel. The Establishment History Panel (*Betriebshistorikpanel, BHP*) is a panel data set that contains information on the number of employees, sector and location of all establishments with at least one dependent employee on 30 June of each year since 1975. Following the regulations of the German Federal Employment Agency, an establishment is defined as the aggregation of all employees in a municipality that are working for the same firm in the same sector. That is, if a firm has several plants in a municipality, all plants in the same sector are assigned the same establishment identifier. Plants in different sectors have distinct identifiers even within the same municipality. Sectors are defined based on the Classification of Economic Activities of the German Statistical Office. The location of establishments is provided at the county level. Germany is divided into 402 counties with around 200,000 inhabitants on average. German counties are roughly comparable to counties in the US. Schmucker et al. (2016) provide a detailed description of the data set.

Extension files on entries and exits of establishments. The establishment identifier in the Establishment History Panel may change when a firm restructures. The extension files render it possible to follow the establishments nonetheless. The files use information on worker flows to identify establishment openings and closings. Hethey and Schmieder (2010) provide details on the files.

A.1.2. Orbis

We use firm-level balance sheet information from the database Orbis of the commercial data provider Bureau van Dijk (BvD). BvD compiles data from publicly available sources as well as by acquiring data from other commercial data providers. For Germany, BvD's main data provider is Creditreform. BvD defines a firm as an independent unit that holds a specific legal form and may incorporate one or more establishments.

BvD's financial information on firms in Germany is most reliable since 2006, as there were changes in the financial reporting system in Germany in that year. In earlier years, a higher share of financial information is missing.

A.1.3. Record linkage procedure

We use a linkage table between the social security records and the Orbis database. The record linkage was performed independently of our project by the German Record Linkage Center (GRLC, see Antoni and Schnell 2019 or www.record-linkage.de for more details). The basis of the linkage was an extract of Orbis acquired by the Institute for Employment Research (IAB). This extract contained data on all German firms

[17] Antoni et al. (2016) focus on the Sample of Integrated Labor Market Biographies (SIAB), a 2% random sample drawn from the IEB.

at the reference date of January 30, 2014. Of the 1,938,990 firms contained in the data, 1,627,668 were marked as active in Germany.

Apart from a wide range of financial variables, the extract contained the name, legal form and address of each firm. The GRLC used these identifiers to link the firm-level data to the administrative establishment-level data of the IAB. This was made possible by the fact that firms have to apply for an establishment number to be issued centrally by the Federal Employment Agency (BA) for each establishment they set up. During this process, firms are required by law to provide their name and legal form as well as the address of the establishment to be recorded in the Data Warehouse (DWH) of the BA. At the time of the record linkage, the DWH included names, the superordinate firm's legal form and addresses of establishments that had been active before or in 2013. To increase the linkage success while also limiting the computational and memory requirements, the GRLC used linkage identifiers of all establishments that had been recorded as active in Germany at least one day during the years 2011 to 2013. Despite this restriction, names, legal forms and addresses of more than 12 million different establishment numbers could be used for the record linkage.

The whole set of identifiers is used to identify the headquarters establishment of the firm. Other establishments within the same firm do not have to be located in the same municipality as the headquarters, which is why additional establishments were linked using only the name and legal form of the firm. In some steps of the iterative linkage process, the GRLC also used the main sector of activity, as this is also contained in both databases.

As these identifiers are non-unique and error-prone, the GRLC developed extensive cleaning, standardization and parsing routines (usually referred to as pre-processing) to achieve records that could successfully be compared between the two data sources. To deal with remaining differences in, for instance, the spelling or abbreviations in identifiers, the GRLC applied error-tolerant methods of record linkage (see Christen 2012). The resulting linkage process consists of 17 consecutive steps, not counting the pre-processing, that varied in terms of which identifiers were used and how strict the requirements on agreement of the compared records were. Schild (2016) provides a more detailed description of the record linkage process. Antoni et al. (2018) report on the linkage success and the representativeness of the resulting data set.

To rule out that we classify independent firms with similar names as ME firms by accident, we only keep establishments that were matched based on the following criteria: exact long name and legal form, exact short name and legal form, exact long name (with or without activity component) and zip code, exact short name (with or without activity component) and zip code.

A.1.4. Identification of headquarters

The record linkage procedure aimed at identifying as many establishments per firm as possible without determining the headquarters of the firm. This information was added by the Research Data Centre (FDZ) at the IAB afterwards. To do so, the FDZ performed several iterative steps that mainly relied on the address of the firm according to Orbis and of the establishments according to the administrative data. During later steps the FDZ also used information on the share of administrative staff or the industry code of the establishments. Antoni et al. (2018) provide details on the process.

A.2. Sector and occupation classification

We use information on the establishment sector at the three digit level. The sector information is based on the respective latest sector classification of the German Federal Statistical Office that updated the classification in 1993, 2003 and 2008. We use the 2008 classification for the 2012 cross-section. We follow Eberle et al. (2011) and transfer the sector classification after 2003 into the classification as of 1993 for the analyses using the 2000-2010 data. Results in the 2012 cross-section are similar if we use the 1993 classification.

The information on the occupation of employees follows the German classification of occupations *"Klassifikation der Berufe"* (KldB). The years 2000-2010 contain the three digit occupation according to the 1988 version of the KldB. The year 2012 contains the five digit occupation according to the 2010 version of the KldB. In 2011, establishments were free to report using either version of the KldB, so we exclude 2011 from our analysis.

A.3. Assignment of occupations to layers

Layers. To assign occupations to layers, we build on the classification of Caliendo et al. (2015) for the French PCS ESE occupation classification. We transfer the classification to the international ISCO classification of occupations and from there to the German occupation classification KldB. Tåg (2013) and Friedrich (2020) use an analogous procedure for Swedish and Danish data. We use official correspondence tables from the German Federal Employment Agency and the International Labor Organization (ILO). In some cases, the translation assigns several layers to the same occupation. Following Friedrich (2020), we generally assign the minimum level layer to these occupations. For the translation between the new and the old German classification KldB 2010 and KldB 1988, we additionally exploit information on the most common correspondence ("*Schwerpunktumsteiger*") to resolve multiple assignments. Table A.1 displays our assignment of occupations to layers. We treat the lowest level layer in each firm as non-managerial.

Blossfeld occupational categories. According to Blossfeld (1983, 208), managers are employees in occupations that have decision-making power over the use of production factors as well as high-level officials in organizations. The assignment from Blossfeld (1983, 1987) treats the following occupations as managerial: 751, 752, 753, 761, 762, 763.

Table A.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
3	751	62194, 63124, 63194, 71104, 73294, 84394, 94494	Manager, executive, director, board member
2	629, 721, 722, 724, 752, 753, 761, 762, 763, 843	All sub-groups of type 2 in occupation groups: 434, 524 434, 524, 815; of type 3 in occupation groups: 411, 431, 434, 524, 922; of type 4 in occupation groups: 411, 412, 431, 432, 433, 434, 511, 513, 516, 524, 532, 632, 633, 712, 713, 715, 722, 731, 814, 815, 921, 922, 933; plus: 11494, 11594, 21194, 23294, 27194, 27294, 27394, 28194, 28294, 28394, 29194, 29294, 31174, 31194, 41203, 41303, 41304, 41383, 41384, 41394, 41403, 41404, 41484, 41494, 42124, 42144, 42314, 42324, 42394, 43152, 43323, 43343, 43353, 43383, 51133, 51233, 51533, 51543, 51594, 53184, 53194, 53394, 61194, 61294, 61394, 62514, 63114, 63313, 63404, 71333, 71433,72144, 72184, 72194, 72243, 72304, 73204, 73214, 73224, 73234, 73244, 73394, 81214, 81234, 81804, 81814, 81884, 82594, 83193, 83194, 84194, 84294, 84304, 84494, 91344, 91354, 92113, 92304, 92394, 92424, 92434, 93303, 93313, 93323, 93343, 93383, 94214, 94404, 94414, 94484, 94493, 94534, 94794	Manager in business organization and strategy, finanical analyst, software developer, qualified IT-specialist, lawyers
1	31, 32, 601, 602, 603, 604, 605, 606, 607, 611, 612, 621, 622, 623, 624, 625, 626, 627, 628, 633, 687, 811, 812, 813, 822, 831, 841, 842, 844, 851, 852, 853, 855, 862, 863, 871, 872, 873, 874, 875, 881, 882, 883, 891, 892, 893, 922	All sub-groups of type 2 in occupation groups: 271, 273, 311, 312, 412, 414, 421, 422, 613, 634, 811, 812, 817, 818, 821, 822, 833, 931, 932, 944, 946, 947; of type 3 in occupation groups: 114, 233, 271, 292, 312, 341, 421, 422, 423, 432, 523, 531, 532, 533, 541, 611, 612, 613, 634, 721, 733, 811, 812, 816, 817, 818, 821, 822, 842, 844, 845, 913, 923, 924, 931, 941, 942, 943, 945, 946, 947; of type 4 in occupation groups: 117, 221, 222, 223, 231, 232, 233, 234, 241, 242, 243, 244, 245, 251, 252, 261, 262, 263, 292, 312, 321, 322, 341, 342, 343, 422, 512, 523, 714, 813, 816, 817, 821, 822, 833, 845, 911, 912, 914, 931, 932, 941, 943, 946; plus: 1104, 11103, 11104, 11113, 11114, 11123, 11124, 11132, 11133, 11183, 11184, 11214, 11233, 11423, 11424, 11603, 11604, 11713, 11723, 12103, 12104, 12113, 12123, 12144, 21113, 21114, 21124, 21213, 21223, 21233, 21313, 21323,21363, 21413, 21423, 22103, 22183, 22203, 22222, 22303, 22333, 22343, 23113, 23123, 23222, 23223, 23322,	Quality manager, training supervisor, management assistant, scientist, engineer, interpreter

Continued on next page

Table A.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
		23413, 23423, 24133, 24203, 24233, 24303, 24413, 24423, 24513, 24523, 24533, 25103, 25133, 25183, 25213, 25223, 25233, 25243, 25253, 26113, 26123, 26223, 26234, 26243, 26253, 26263, 26303, 26313, 26323, 26333, 26383, 27104, 27184, 27212, 27223, 27224, 27283, 27284, 27303, 27304, 27313, 27314, 28103, 28104, 28113, 28114, 28123, 28133, 28143, 28213, 28214, 28223, 28224, 28313, 28314, 28333, 28343, 29103, 29104, 29113, 29114, 29123, 29133, 29134, 29143, 31103, 31104, 31114, 31124, 31133, 31134, 31143, 31144, 31153, 31154, 31163, 31164, 31173, 32103, 32113, 32123, 32203, 32223, 32233, 32243, 32253, 32263, 33133, 33213, 33223, 33233, 33243, 33303, 33323, 34203, 34213, 34233, 34303, 34323, 34343, 41213, 41283, 41293, 41313, 41314, 41322, 41323, 41324, 41413, 41414, 41423, 41424, 41433, 41434, 41483, 42114, 42134, 42334, 43102, 43112, 43122, 43313, 43333, 43363, 51113, 51123, 51182, 51183, 51223, 51243, 51503, 51504, 51513, 51523, 51534, 51583, 51593, 51623, 51633, 51643, 51653, 51663, 53124, 53134, 53152, 53222, 53232, 53312, 53314, 53322, 53332, 61124, 61132, 61204, 61214, 61284, 61314, 62183, 62513, 63122, 63123, 63132, 63212, 63213, 71303, 71383, 71403, 71423, 71522, 71523, 72124, 72134, 72213, 72223, 72233, 72303, 73162, 73163, 73183, 73202, 73203, 73212, 73213, 73232, 73233, 73241, 73242, 73243, 73253, 73254, 73282, 73283, 73284, 73314, 73324, 73334, 81224, 81294, 81302, 81313, 81323, 81332, 81333, 81352, 81353, 81382, 81383, 81393, 81894, 82332, 82343, 82494, 82503, 82504, 82514, 82522, 82523, 82524, 82534, 83112, 83123, 83124, 83132, 83133, 83134, 83154, 83223, 83323, 83333, 83383, 84114, 84124, 84134, 84144, 84184, 84214, 84224, 84404, 84412, 84414, 84424, 84434, 84444, 84454, 84484, 91233, 91314, 91324, 91334, 91384, 92133, 92384, 92414, 92494, 93213, 93223, 93233, 93383, 93413, 93433, 93513, 93523, 93524, 93603, 93604, 93613, 93623, 93633, 93643, 93653, 93683, 94224, 94403, 94413, 94483, 94514, 94522, 94532, 94582, 94704, 94714, 94724	
0	Others	Others	Unskilled/semi-skilled occupations in metal-working, printing, machine and equipment assemblers, green keepers, catering, office clerks

The KldB 1988 assigns a three digit code to each occupation. The KldB 2010 assigns a five digit code to each occupation. The first three digits denote the occupation group. Digit 4 denotes the occupation sub-group. Digit 5 denotes the type of occupation (1 = unskilled/semi-skilled, 2 = skilled, 3 = complex, 4 = highly complex).

A.4. Plausibility of the assignment of occupations to layers

A.4.1. Evidence for full sample based on Caliendo et al. (2015)

Firms form a hierarchy. Table A.2 presents summary statistics of the wage distributions by layer in the 2000-2010 data and the 2012 cross-section. We provide the same statistics on education, as this information is observable in our data. The higher ends of the wage distributions show the same values because the daily wages are censored at the social security limit. Both wages and education are higher for higher layers.

Table A.2: Distribution of daily wages and education by layer

(a) Wages, 2000-2010 data

	Layer 0	Layer 1	Layer 2	Layer 3
mean	98.02	138.27	144.92	155.13
p5	46.04	74.91	75.26	84.18
p10	55.35	86.92	93.69	104.50
p25	73.08	109.37	123.75	145.79
p50	93.69	147.61	155.10	168.23
p75	119.19	172.14	176.38	176.38
p90	151.18	176.38	176.38	176.38
p95	168.23	176.38	176.38	176.38
N	43885369	8750156	1274808	1075692

(b) Wages, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	91.47	124.22	147.15	151.66
p5	43.54	59.71	78.55	72.18
p10	50.47	71.23	93.85	90.29
p25	65.33	92.47	124.01	135.43
p50	85.69	122.80	159.15	176.38
p75	111.29	163.52	176.38	176.38
p90	144.73	176.38	176.38	176.38
p95	167.71	176.38	176.38	176.38
N	4271175	1378410	629518	76969

(c) Education, 2000-2010 data

	Layer 0	Layer 1	Layer 2	Layer 3
mean	1.99	2.92	2.80	3.06
p5	1	2	2	2
p10	1	2	2	2
p25	2	2	2	2
p50	2	3	2	4
p75	2	4	4	4
p90	3	4	4	4
p95	4	4	4	4
N	43885369	8750156	1274808	1075692

(d) Education, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	2.00	2.74	3.23	3.01
p5	1	2	2	2
p10	1	2	2	2
p25	2	2	2	2
p50	2	2	4	3
p75	2	4	4	4
p90	3	4	4	4
p95	4	4	4	4
N	4271175	1378410	629518	76969

This table is based on Table 1 in Caliendo et al. (2015). It reports mean and percentiles of the daily wage distribution in 2010 Euros and the education distribution for each layer. Panels (a) and (c) contain all firm-years in the 2000-2010 data. Panels (b) and (d) contain the firms in the 2012 cross-section. We separate the years 2000-2010 and 2012 because layers are assigned based on the classification of occupations that differs between these years (see Appendix A.2). The 90th and 95th percentiles of wages for layers 1-3 and the 75th percentile for layers 2-3 are equal because daily wages are censored at the social security limit. Education levels: 1 - Primary school/ lower secondary school/ intermediate school leaving certificate, no vocational qualification; 2 - As 1, but with vocational qualification; 3 - Upper secondary school leaving certificate (Abitur); 4 - Degree from university/ university of applied sciences.

Table A.3 documents that firm characteristics differ systematically depending on the number of managerial layers.

Table A.3: Data description by number of managerial layers

	2000-2010 data				2012 cross-section			
		— Average —		Median		— Average —		Median
# lyrs	firm-yrs	# emp	wage	wage	firm-yrs	# emp.	wage	wage
0	189,453	26	78.88	79.61	22,068	19	73.73	73.89
1	237,767	37	88.07	88.04	32,573	25	81.38	80.06
2	162,860	66	96.97	96.73	34,130	54	94.07	92.55
3	106,900	287	108.19	108.33	20,586	159	106.70	106.22

This table is based on Table 3 in Caliendo et al. (2015). It reports descriptive statistics for firms by the number of managerial layers. We separately present statistics for the years 2000-2010 and 2012 because layers are assigned based on the classification of occupations that differs between these years (see Appendix A.2). *# of emp.* refers to the number of employees. *wage* refers to daily wages.

Tables A.4 to A.6 present evidence that firms form a hierarchy. Table A.4 presents the share of firms with consecutively ordered layers. The shares are similar to Caliendo et al. (2015) for the 2012 cross-section. In the 2000-2010 data, the share is lower for firms with two managerial layers. This is driven by the fact that the top managerial occupation, 751, is very broad in the classification of occupations used in the 2000-2010 data, so about half of firms with two managerial layers have layers 0, 1, and 3.

Table A.4: Percentage of firms that have consecutively ordered layers

(a) 2000-2010 data

	Among firms with				
	0 layers	1 layer	2 layers	3 layers	All Firms
Unweighted	99.81	63.81	30.87	100.00	71.45
Weighted by # emp.	99.86	63.89	37.50	100.00	82.00

(b) 2012 cross-section

	Among firms with				
	0 layers	1 layer	2 layers	3 layers	All Firms
Unweighted	97.20	69.15	76.55	100.00	82.93
Weighted by # emp.	97.44	70.08	86.32	100.00	92.04

This table is based on Table 4 in Caliendo et al. (2015). It reports the fraction of firms with consecutively ordered layers for different numbers of managerial layers. The first row presents the unweighted share. The second row presents the shares weighted by the total number of employees. We separately present statistics for the years 2000-2010 and 2012 because layers are assigned based on the classification of occupations that differs between these years (see Appendix A.2).

Table A.5 shows that the share of firms that satisfy a hierarchy in the number of employees per layer tends to be higher than the share of firms that satisfy a hierarchy in hours in Caliendo et al. (2015). The share of firms that satisfy a hierarchy in daily wages is similar to the share of firms that satisfy a hierarchy in hourly wages in Caliendo et al. (2015) according to Table A.6.

Table A.5: Firms that satisfy a hierarchy in the # of employees

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# lyrs	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$
1	93.8	93.8			89.0	89.0		
2	76.9	91.9	84.9		65.0	80.6	81.5	
3	47.6	91.1	88.6	62.6	50.9	75.7	76.9	91.9

This table is based on Table 5 in Caliendo et al. (2015). It reports the share of firms that satisfy a hierarchy in the number of employees separately for firms with 1, 2, and 3 managerial layers. Columns (1) and (5) report the fraction of firms that satisfy a hierarchy at all layers. Columns (2)-(4) and (6)-(8) display the fraction of firms that satisfy a hierarchy at layers 1, 2, and 3, respectively. A firm satisfies a hierarchy in the number of employees if the number of employees at layer $\ell - 1$ exceeds the number of employees at layer ℓ.

Table A.6: Firms that satisfy a hierarchy in wages

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# lyrs	$w_L^{\ell+1} \geq w_L^\ell$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^{\ell+1} \geq w_L^\ell$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	90.4	90.4			84.3	84.3		
2	64.9	90.9	73.4		66.4	84.5	80.8	
3	37.5	94.2	57.8	80.3	61.8	86.6	82.8	88.9

This table is based on Table 6 in Caliendo et al. (2015). It reports the share of firms that satisfy a hierarchy in wages separately for firms with 1, 2, and 3 managerial layers. Columns (1) and (5) report the fraction of firms that satisfy a hierarchy at all layers. Columns (2)-(4) and (6)-(8) display the fraction of firms that satisfy a hierarchy at layers 1, 2, and 3, respectively. A firm satisfies a hierarchy in wages if wages at layer ℓ exceed the wages at layer $\ell - 1$.

How do firms expand? Table A.7 documents that firm level outcomes change systematically depending on whether firms add, keep constant or drop managerial layers. Firms that add layers increase their number of employees, but decrease wages at preexisting layers. Firms that drop layers decrease their number of employees, but increase wages at preexisting layers.

Table A.7: Change in firm-level outcomes, 2000-2010 data

		All	Increase L	No Change in L	Decrease L
d	ln # employees	0.008***	0.089***	0.007***	-0.059***
	Detrended		0.080***	-0.001***	-0.066***
d	ln # emp., normalized	0.009***	1.485***	0.004***	-1.521***
	Detrended		1.476***	-0.005***	-1.529***
d	ln average wage	-0.003***	0.006***	-0.003***	-0.014***
	Detrended		0.009***	0.000*	-0.011***
	Common layers	0.001***	-0.015***	-0.003***	0.087***
	Detrended		-0.017***	-0.004***	0.085***

This table is based on Table 11 in Caliendo et al. (2015). It reports changes (d) in firm-level outcomes between two consecutive years for all firms, firms that increase, keep constant, and decrease managerial layers. We detrend variables by their respective annual averages. *** p < .01, * p < .1.

Tables A.8-A.10 document how the number of employees, wages and measures of knowledge change at all layers if sales change, but firms do not change their number of layers. Consistent with the results in Caliendo et al. (2015), wages are positively related to sales at all layers. The low coefficient for layers 2 and 3 likely reflects that wages are censored at the social security limit for more than a quarter of observations in those layers (see Table A.2). The knowledge of employees, as reflected by their experience, significantly increases at higher layers, similar to Caliendo et al. (2015). Unlike in that paper, formal education does not change significantly. This may reflect the coarseness of our measure with only four values.

Table A.8: Elasticity of # of employees with sales for firms that do not change L, 2000-2010 data

# of layers	Layer	β_L^l	Standard Error	p-Value	# obs.
1	0	0.046	0.008	0.000	41,415
2	0	0.058	0.008	0.000	27,153
2	1	0.042	0.009	0.000	27,153
3	0	0.036	0.008	0.000	24,626
3	1	0.011	0.008	0.157	24,626
3	2	0.003	0.007	0.711	24,626

This table is based on Table 9 in Caliendo et al. (2015). It reports the results of regressions of detrended log change in normalized number of employees at layer l in a firm with L layers on its detrended log change in sales, and no constant, selecting all the firms that stay at L layers across two consecutive years. The term β_L^l is the coefficient on log change in sales.

Table A.9: Elasticity of wages with sales for firms that do not change L, 2000-2010 data

# of layers	Layer	β	Standard Error	p-Value	# obs.
0	0	0.015	0.002	0.000	36,144
1	0	0.015	0.001	0.000	41,415
1	1	0.016	0.002	0.000	41,415
2	0	0.016	0.002	0.000	27,153
2	1	0.011	0.003	0.000	27,153
2	2	0.014	0.002	0.000	27,153
3	0	0.013	0.001	0.000	24,626
3	1	0.010	0.002	0.000	24,626
3	2	0.011	0.002	0.000	24,626
3	3	0.002	0.001	0.096	24,626

This table is based on Table 10 in Caliendo et al. (2015). It reports the results of regressions of the log change in daily wages by layer on the log change in sales for firms that do not change their number of layers L across two periods. We detrend both variables by their respective annual averages.

Table A.10: Elasticity of knowledge with sales for firms that do not change L, 2000-2010 data

# of layers	Layer	Experience	p-Value	Education	p-Value	# obs.
0	0	0.002	0.758	0.001	0.731	36,144
1	0	−0.004	0.383	−0.001	0.443	41,392
1	1	0.034	0.000	0.001	0.737	40,921
2	0	−0.012	0.039	−0.001	0.255	27,152
2	1	0.002	0.789	0.000	0.917	26,923
2	2	0.027	0.001	0.001	0.540	26,716
3	0	−0.017	0.000	−0.002	0.041	24,625
3	1	−0.005	0.258	−0.001	0.541	24,530
3	2	0.011	0.079	0.003	0.112	24,450
3	3	−0.003	0.614	0.001	0.687	24,471

This table is based on Table 15 in Caliendo et al. (2015). It reports the results of regressions of the log change in days of experience and of education by layer on the log change in sales for firms that do not change their number of layers L across two periods. We detrend both variables by their respective annual averages.

Tables A.11-A.14 document how the number of employees, wages and measures of knowledge change at all layers if sales change and firms adjust their number of layers. The number of employees increase (decrease) at preexisting layers if firms add (drop) a layer; coefficient signs and significance levels in Table A.11 are very similar to those in Caliendo et al. (2015). The relative size of coefficients is also similar (although we do not test whether differences are significant): for example, the increase in the number of employees is the higher, the more managerial layers firms without management add, just as the increase in the number of hours in Caliendo et al. (2015). Wages decrease (increase) at preexisting layers if firms add (drop) a layer; coefficient signs, significance levels, and the relative coefficient size in Table A.12 are again very similar to those in Caliendo et al. (2015). After adding layers, only the new managers earn more according to Table A.13, which is again consistent with Caliendo et al. (2015). The results for the knowledge of employees as reflected by their experience in Table A.14 are similar to Caliendo et al. (2015). The results for formal education are more significant than in Table A.10 and generally in line with Caliendo et al. (2015).

Table A.11: Average log change in # of employees for firms that change L, 2000-2010 data

# of layers before	# of layers after	Layer	$d \ln \tilde{n}^{\ell}_{Lit}$	Standard Error	p-Value	# obs.
0	1	0	3.059	0.008	0.000	12,168
0	2	0	3.200	0.029	0.000	828
0	3	0	3.362	0.140	0.000	63
1	0	0	−3.072	0.007	0.000	11,820
1	2	0	0.794	0.009	0.000	14,542
1	2	1	0.659	0.011	0.000	14,542
1	3	0	1.083	0.046	0.000	900
1	3	1	0.880	0.049	0.000	900
2	0	0	−3.231	0.032	0.000	631
2	1	0	−0.798	0.010	0.000	13,663
2	1	1	−0.674	0.011	0.000	13,663
2	3	0	0.205	0.007	0.000	9,828
2	3	1	0.136	0.008	0.000	9,828
2	3	2	−0.106	0.010	0.000	9,828
3	0	0	−3.160	0.126	0.000	44
3	1	0	−1.010	0.055	0.000	597
3	1	1	−0.843	0.058	0.000	597
3	2	0	−0.269	0.008	0.000	8,561
3	2	1	−0.181	0.009	0.000	8,561
3	2	2	0.039	0.011	0.000	8,561

This table is based on Table 12 in Caliendo et al. (2015). It reports estimates of the average detrended log change in the normalized number of employees $d \ln \tilde{n}^{\ell}_{Lit}$ at each layer among firms that change their number of layers. We detrend the number of employees by its annual average.

Table A.12: Average log change in wages for firms that change L, 2000-2010 data

# of layers before	# of layers after	Layer	$d\ln\tilde{w}^{\ell}_{Lit}$	Standard Error	p-Value	# obs.
0	1	0	-0.010	0.001	0.000	12,168
0	2	0	-0.012	0.003	0.000	828
0	3	0	-0.030	0.016	0.067	63
1	0	0	0.004	0.001	0.000	11,820
1	2	0	-0.005	0.001	0.000	14,542
1	2	1	-0.101	0.003	0.000	14,542
1	3	0	-0.001	0.003	0.828	900
1	3	1	-0.101	0.003	0.000	900
2	0	0	0.009	0.003	0.007	631
2	1	0	0.004	0.001	0.000	13,663
2	1	1	0.083	0.003	0.000	13,663
2	3	0	0.000	0.001	0.872	9,828
2	3	1	-0.021	0.002	0.000	9,828
2	3	2	-0.142	0.003	0.000	9,828
3	0	0	0.063	0.024	0.013	44
3	1	0	0.012	0.003	0.001	597
3	1	1	0.087	0.013	0.000	597
3	2	0	0.006	0.001	0.000	8,561
3	2	1	0.017	0.002	0.000	8,561
3	2	2	0.112	0.003	0.000	8,561

This table is based on Table 13 in Caliendo et al. (2015). It reports estimates of the average detrended log change in the daily wages $d\ln\tilde{w}^{\ell}_{Lit}$ at each layer among firms that change their number of layers. We detrend wages by their annual average.

Table A.13: Decomposition of total log change in average wages, 2000-2010 data

From/ To	Relative average wage, common layers 1	2	3	From/ To	Relative average wage, added layers 1	2	3
0	0.993***	0.992***	0.978	0	1.385***	1.416***	1.390***
	(12,155)	(828)	(63)		(12,156)	(828)	(63)
1		0.988***	0.990***	1		1.491***	1.413***
		(14,527)	(899)			(14,527)	(900)
2			0.989***	2			1.517***
			(9,819)				(9,818)

From/ To	Share of employees, common layers 1	2	3	From/ To	Overall mean change in wages 1	2	3
0	0.943***	0.898***	0.866***	0	0.010***	0.023***	0.008
	(12,156)	(828)	(63)		(12,156)	(828)	(63)
1		0.950***	0.912***	1		0.008***	0.016***
		(14,528)	(899)			(14,528)	(900)
2			0.956***	2			0.006***
			(9,820)				(9,819)

This table is based on Table 14 in Caliendo et al. (2015). *** $p < .01$ It reports the sources of change in the daily wages by type of transition. The upper left panel reports average daily wages after transition relative to those before transition in common layers. The upper right panel reports average daily wages after transition relative to those before transition in the newly added layers. Each cell is computed excluding observations below the 0.05th and above the 99.95th percentile. Numbers of observations are in parentheses.

Table A.14: Average change in knowledge for firms that change L, 2000-2010 data

# lyrs before	# lyrs after	Layer	Experience	p-Value	Education	p-Value	# obs.
0	1	0	0.001	0.679	−0.003	0.000	12,154
0	2	0	−0.017	0.123	−0.007	0.024	828
0	3	0	−0.077	0.265	0.001	0.968	63
1	0	0	0.006	0.003	0.003	0.000	11,813
1	2	0	−0.007	0.000	−0.001	0.045	14,537
1	2	1	−0.469	0.000	0.005	0.033	14,165
1	3	0	−0.051	0.000	−0.005	0.162	900
1	3	1	−0.406	0.000	−0.002	0.872	868
2	0	0	0.005	0.660	0.010	0.005	631
2	1	0	−0.000	0.862	0.003	0.000	13,659
2	1	1	0.193	0.000	0.004	0.090	13,354
2	3	0	−0.013	0.000	0.001	0.126	9,825
2	3	1	−0.159	0.000	0.006	0.004	9,666
2	3	2	−0.656	0.000	−0.027	0.000	9,517
3	0	0	0.022	0.652	−0.009	0.694	44
3	1	0	0.001	0.923	0.011	0.018	597
3	1	1	0.196	0.000	0.009	0.466	581
3	2	0	0.002	0.507	0.005	0.000	8,561
3	2	1	0.002	0.000	−0.005	0.019	8,439
3	2	2	0.196	0.000	0.046	0.000	8,322

This table is based on Table 16 in Caliendo et al. (2015). It reports estimates of the average detrended log change in days of experience and of education at each layer for firms that change their number of layers. We detrend both variables by their respective annual averages.

A.4.2. Evidence on the tasks of occupations by layer

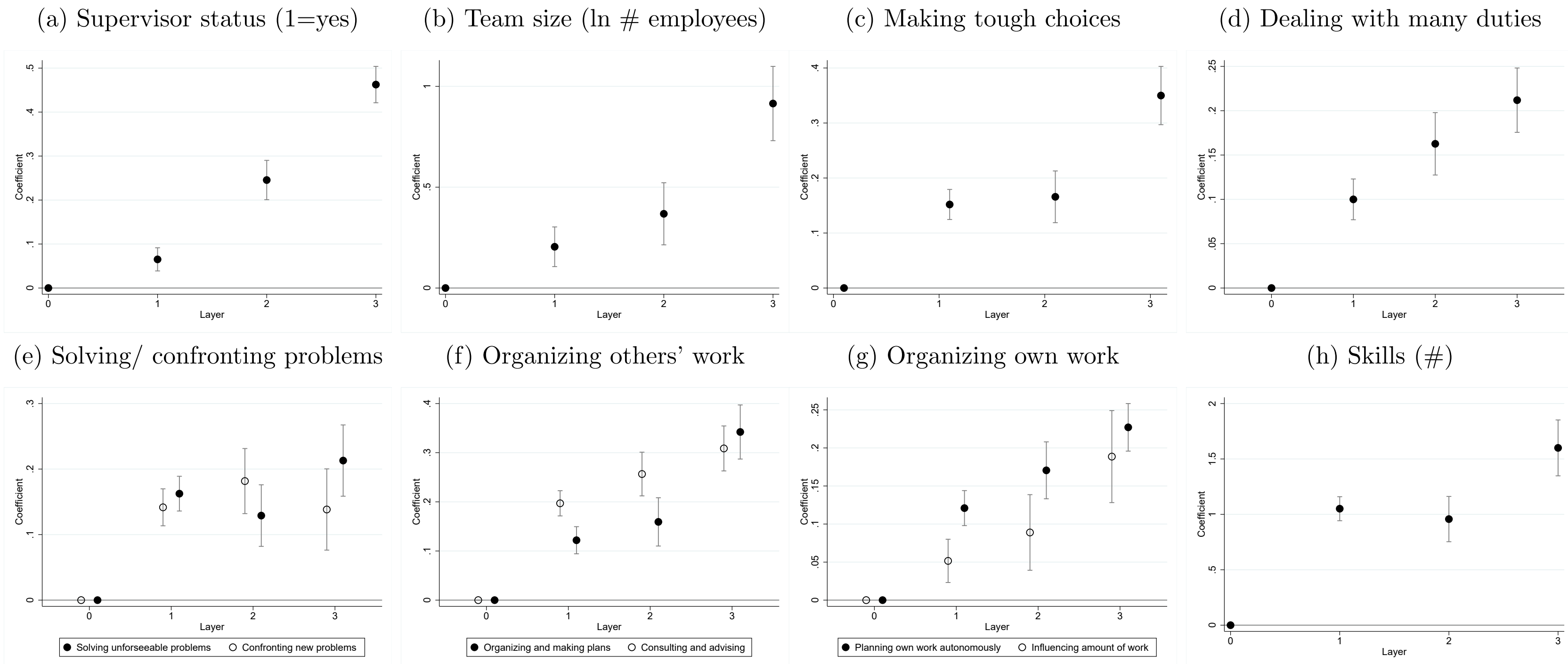
The *2006 BIBB/BAuA Survey of the Working Population* administered by the German Federal Institute for Vocational Education and Training (*Bundesinistitut für Berufsbildung, BIBB*) and the Federal Institute for Occupational Safety and Health (*Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, BAuA*) collects data on the education, career and current employment conditions of a representative sample of 20,000 working age individuals in Germany (Hall and Tiemann 2006). The data contain information on the occupation of employees (KldB 1988 classification). We relate the tasks of employees to the layer assigned to their occupation by estimating, via OLS:

$$y_i = \beta \mathbf{D}_{\text{layer},i} + \gamma \mathbf{X}_i + \delta \mathbf{Z}_i + u_i \tag{A.4.1}$$

where y_i is individual i's answer to a survey question about i's tasks, $\mathbf{D}_{\text{layer},i}$ is a dummy for the layer to which we assign individual i's occupation, $\mathbf{X}_i$ is a vector of employee characteristics and $\mathbf{Z}_i$ are characteristics of i's employer.

Figure A.1 plots the coefficients and 95% confidence bands by layer. Table A.15 presents the regression results.

Figure A.1: Evidence on tasks by layer, 2006 BIBB/BAuA survey



The figure plots the estimated coefficients of the layer dummies in equation (A.4.1). In figures (c)-(g), 1=often. See notes of Table A.15 for the survey questions.

Summary. Employees at higher layers are more likely to be supervisors. The median predicted probability that an employee at layer 3 is a supervisor is 84%. Employees at higher layers supervise larger teams. They are more likely to take decisions, have more responsibilities, solve unforeseeable and confront new problems, and organize work for others. They are more independent in organizing their own work. The job of employees at higher layers also require more skills. Overall, this descriptive evidence corroborates the assumption that the assignment of occupations to layers reflects differences between the managerial tasks and duties of employees in firms.

Table A.15: Regression results: tasks by layer, 2006 BiBB/BAuA survey

	(a)	(b)	(c)	(d)	(e1)	(e2)	(f1)	(f2)	(g1)	(g2)	(h)
Layer 1	0.065***	0.205***	0.152***	0.100***	0.162***	0.142***	0.122***	0.197***	0.121***	0.052***	1.051***
	(0.013)	(0.050)	(0.014)	(0.012)	(0.014)	(0.014)	(0.014)	(0.013)	(0.012)	(0.014)	(0.055)
Layer 2	0.245***	0.368***	0.166***	0.163***	0.129***	0.182***	0.159***	0.257***	0.171***	0.089***	0.957***
	(0.023)	(0.079)	(0.024)	(0.018)	(0.024)	(0.025)	(0.025)	(0.022)	(0.019)	(0.025)	(0.104)
Layer 3	0.463***	0.915***	0.350***	0.212***	0.213***	0.138***	0.342***	0.309***	0.227***	0.189***	1.600***
	(0.021)	(0.094)	(0.027)	(0.019)	(0.028)	(0.031)	(0.028)	(0.023)	(0.016)	(0.030)	(0.129)
Age	0.002***	0.007***	0.001	−0.001**	−0.006***	−0.002***	−0.002***	−0.002***	0.003***	0.005***	−0.009***
	(0.001)	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Tenure	−0.000	−0.001**	−0.000*	−0.000	−0.000**	0.000	−0.000	−0.000	0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gender	−0.122***	−0.193***	−0.095***	0.019	−0.084***	−0.068***	−0.021*	0.095***	0.037***	−0.057***	−0.454***
	(0.010)	(0.041)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)	(0.041)
Constant	0.002	0.645**	0.639***	0.646***	0.800***	0.515***	0.283***	0.496***	0.611***	0.403***	2.764***
	(0.039)	(0.233)	(0.046)	(0.045)	(0.048)	(0.047)	(0.045)	(0.046)	(0.048)	(0.104)	(0.179)
# observations	13,818	4,857	13,824	13,823	13,825	13,824	13,807	13,826	13,272	13,238	13,828
F ($\beta_1 = \beta_2$)	55.01***	3.82+	0.30	11.05***	1.78	2.35	1.98	6.79**	6.48*	2.01	0.74
F ($\beta_2 = \beta_3$)	55.32***	21.37***	29.02***	4.41*	5.84*	1.30	26.00***	3.11+	6.46*	7.05**	16.20***
F ($\beta_1 = \beta_3$)	300.38***	49.76***	49.30***	32.63***	3.07+	0.01	56.41***	21.99***	40.18***	18.87***	16.91***

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* (a) Supervisor status (1=yes); (b) How many people do you supervise? (ln #) (c) How frequently do you make tough choices on your own responsibility? (d) How frequently do you deal with a range of duties and responsibilities? (e1) How frequently do you have to react to and solve unforeseeable problems? (e2) How frequently are you confronted with new problems? (f1) How frequently does the task of organizing and making plans appear in your job? (f2) How frequently does the task of consulting and advising appear in your job? (g1) How frequently are you allowed to plan and schedule your work by yourself? (g2) How frequently are you able to influence the amount of work you have to do? (h) Number of subject areas in which specialized skills are required. For questions (c)-(g), 1=often, 0=sometimes–never. *Independent variables: Layer X:* indicator variable: occupation assigned to layer X; *Age:* age of respondent in years; *Tenure:* tenure of respondent in decades; *Gender:* gender of respondent, 1=female. Education, firm size and sector category fixed effects included. *F ($\beta_j = \beta_k$):* F-statistic, test for equality of coefficients of Layer j and Layer k.

A.5. Descriptive statistics

Table A.16: Share of ME firms by sector, 2012 cross-section

Broad sector	N	% share ME in		
		firms	est.s	empl.
Manufacturing	31,418	7	20	39
Wholesale and retail trade	20,900	12	40	33
Construction	16,163	4	11	13
Services	7,938	14	55	50
Traffic	7,108	8	37	34
Information and Communication	5,660	10	28	33
Other services	4,867	11	35	30
Health	4,353	10	26	20
Hotels and restaurants	2,908	8	33	21
Agriculture	1,479	6	12	11
Other serivces	1,355	10	42	34
Water/Sewage	1,244	7	18	19
Real estate	1,168	11	32	22
Electricity	822	11	48	34
Art and entertainment	775	9	31	12
Finance	717	16	58	42
Mining and quarrying	352	16	38	60

Descriptive statistics on number of firms, share of ME firms in number of firms, establishments (*est.s*) and employees (*empl.*) by sector. Number of observations is lower than total number of firms, because statistics for firms in small sectors were not disclosed for confidentiality.

Note: Table A.16 uses broad sector groups for confidentiality. The regressions in Tables III-V and the corresponding robustness checks include 3-digit sector dummies.

A.5.1. Descriptive statistics for the 2000–2010 panel

Table A.17: Descriptive statistics, SE vs. ME firms, 2000–2010 data

Units of observation	N	of which ME firms (% share)						
Firms	985,914	9.7						
with non-missing sales	332,723	10.7						
Establishments	1,314,254	32.2						
Employees	60,291,082	37.4						
Descriptive statistics	**N**	**ME**	**Mean**	**SD**	**p25**	**p50**	**p75**	**p95**
# employees	890,578	0	42	161	13	20	38	131
	95,336	1	237	2,210	22	48	126	678
Sales (M €)	297,296	0	34	773	2	4	9	71
	35,427	1	362	3,658	4	17	80	716
# managerial layers	890,578	0	1.1	1.0	0	1	2	3
	95,336	1	1.6	1.1	1	2	3	3
Managerial share	890,578	0	16	20	0	9	22	64
(%, layers)	95,336	1	20	22	3	13	30	72
Managerial share	890,578	0	5	9	0	0	6	20
(%, Blossfeld)	95,336	1	6	11	0	2	7	23

Descriptive statistics. *Variable definitions:* see Table I.

Table A.18: Descriptive statistics, ME firms, 2000–2010 data

Descriptive statistics, firm	N		Mean	SD	p50	p75	p95
# establishments (incl. HQ)	95,336		4.4	30.7	2	3	9
Maximum distance to HQ, km	95,336		212	185	164	365	536
Minimum area covered, km^2	31,323		27,931	39,626	6,817	44,454	117,952
Descriptive statistics, est.	**N**	**HQ**	**Mean**	**SD**	**p25**	**p50**	**p75**
# employees	328,346	0	37	368	2	6	17
	95,330	1	109	597	11	28	78
# managerial layers	328,346	0	0.6	0.8	0	0	1
	95,330	1	1.4	1.1	0	1	2
Managerial share	328,346	0	18	29	0	0	25
(%, layers)	95,330	1	21	24	0	13	32
Managerial share	328,346	0	5	15	0	0	0
(%, Blossfeld)	95,330	1	7	14	0	1	8

Descriptive statistics, ME firms. *Variable definitions:* see Table II.

A.5.2. ME firm organization vs. SE firm organization

Tables A.19 and A.20 document that ME firms have more managerial layers than SE firms, both in the 2012 cross-section and the 2000-2010 data.

Table A.19: Managerial organization, ME vs. SE firms, 2012 cross-section

Dependent variable	# managerial layers			Mg. share $\in [0,1]$	
				Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)
ME firm (0/1)	0.262***	0.064***	0.141***	0.229***	0.155***
	(0.007)	(0.008)	(0.005)	(0.013)	(0.017)
Log sales		0.179***			
		(0.005)			
Log # non-mg. employees			0.144***		
			(0.008)		
# firms	105,957	53,568	105,957	105,957	105,957
HQ sector FE	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y
Model		Poisson		GLM	

The table presents the coefficient estimates. Constant included. Standard errors clustered by headquarter county in parentheses. *** $p < 0.001$. *Dependent variable:* (1)-(3) number of managerial layers, (4),(5) managerial share in the wage sum. *Independent variables: ME firm (0/1):* indicator variable for ME firm status, others see Table IV.

Table A.20: Managerial organization, ME vs. SE firms, 2000-2010 data

Dependent variable	# managerial layers			Mg. share $\in [0,1]$	
				Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)
ME firm (0/1)	0.322***	0.086***	0.081***	0.225***	0.103***
	(0.007)	(0.008)	(0.007)	(0.016)	(0.020)
Log sales		0.175***			
		(0.005)			
Log # non-mg. employees			0.262***		
			(0.007)		
# firms	385,529	105,059	385,529	385,529	385,529
HQ sector FE	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Model		Poisson		GLM	

The table presents the coefficient estimates. Constant included. Standard errors clustered by headquarter county in parentheses. *** $p < 0.001$. *Dependent variable:* (1)-(3) number of managerial layers, (4),(5) managerial share in the wage sum. *Independent variables:* see Table A.19.

A.5.3. Within ME firm evidence based on Caliendo et al. (2015)

HQ and establishments form a hierarchy. Table A.21 presents summary statistics of the wage distribution by layer in the 2000-2010 data and the 2012 cross-section separately for headquarters (HQ) and establishments of ME firms. The higher ends of the distributions show the same values as the daily wages are censored at the social security limit. In general, wages increase with higher layers.

Table A.21: Distribution of daily wages by layer

(a) Headquarters, 2000-2010 data

	Layer 0	Layer 1	Layer 2	Layer 3
mean	109.98	147.99	152.44	163.30
p5	53.61	85.34	90.92	105.10
p10	64.63	98.17	109.76	130.75
p25	83.33	126.63	136.80	166.99
p50	106.23	162.38	166.86	176.38
p75	135.33	176.38	176.38	176.38
p90	167.32	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	7621377	1936645	303579	266129

(b) Headquarters, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	105.38	138.59	155.88	163.56
p5	47.65	72.12	91.90	95.33
p10	57.25	85.02	109.28	120.47
p25	76.29	110.06	142.32	167.41
p50	100.82	146.86	176.38	176.38
p75	132.69	176.38	176.38	176.38
p90	166.35	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	602015	286817	147553	15803

(c) Establishments, 2000-2010 data

	Layer 0	Layer 1	Layer 2	Layer 3
mean	108.10	147.95	149.96	159.31
p5	52.86	87.65	88.82	94.53
p10	65.26	100.36	104.72	111.98
p25	83.39	126.53	132.40	151.18
p50	105.62	160.09	163.26	176.38
p75	130.90	176.38	176.38	176.38
p90	159.93	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	9115852	2156689	375875	230552

(d) Establishments, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	107.41	137.50	155.40	127.84
p5	47.99	69.63	92.00	59.58
p10	58.22	82.48	108.73	70.38
p25	77.57	106.64	141.97	92.57
p50	103.82	147.82	176.38	127.18
p75	136.59	176.38	176.38	176.38
p90	167.57	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	742733	257425	111631	12796

This table is based on Table 1 in Caliendo et al. (2015). It reports mean and percentiles of the daily wage distribution for each layer in 2010 Euros, separately for the HQ and establishments of ME firms. We separate the years 2000-2010 and 2012 because of different classifications of occupations (see Appendix A.2). The 95th percentiles, the 75th and 90th percentiles for layers 1-3, and the median for layers 2-3 are equal because daily wages are censored at the social security limit.

An exception are below-median and median wages of employees in layer 3 of establishments in the 2012 cross-section. The wage distribution is also notably different from the distribution in layer 3 for the HQ, unlike the distribution for the other layers. We find that the low-wage employees in layer 3 typically have the occupation 62194 ("managers in retail trade") and work in the retail sector. They work in very small establishments that only have 1.2 full-time employees on average, but may have additional employees who work part-time or on an hourly basis. Table A.22 presents the wage distribution after dropping employees in the retail sector. While wages in the bottom half of the distribution for layer 3 of establishments are still lower than in layer 2, dropping the retail sector reduces the difference by up to 90%. In unreported regressions, we confirm that Fact 2 is robust to dropping the retail sector.

Table A.22: Distribution of daily wages by layer, retail not included

(a) Headquarters, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	106.62	139.04	156.57	167.72
p5	48.40	72.88	93.45	120.00
p10	58.45	85.83	110.85	149.36
p25	77.78	110.83	144.09	176.38
p50	102.22	147.48	176.38	176.38
p75	134.03	176.38	176.38	176.38
p90	167.43	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	579857	283666	143182	13801

(b) Establishments, 2012 cross-section

	Layer 0	Layer 1	Layer 2	Layer 3
mean	109.56	138.04	156.36	154.46
p5	49.46	71.08	94.51	91.76
p10	60.32	83.61	111.34	102.04
p25	80.31	107.55	144.35	139.70
p50	106.50	148.45	176.38	176.38
p75	138.56	176.38	176.38	176.38
p90	169.40	176.38	176.38	176.38
p95	176.38	176.38	176.38	176.38
N	700577	254260	108818	6148

See Table A.21.

Tables A.23 and A.24 present summary statistics by the number of managerial layers. Table A.23 counts the number of managerial layers in an establishment or HQ. Table A.24 reports statistics given the number of managerial layers in the firm.

Table A.23: Data description by number of managerial layers in HQ/establishment

(a) Headquarters

	2000-2010 data				2012 cross-section			
		— Average —		Median		— Average —		Median
# lyrs	HQ-yrs	# emp	wage	wage	HQ-yrs	# emp	wage	wage
0	17,908	23	80.67	79.43	1,633	14	73.00	70.13
1	21,317	40	94.88	92.80	2,161	23	87.52	82.52
2	18,429	88	104.87	103.83	3,104	84	102.14	100.42
3	19,920	363	118.30	117.97	2,914	247	115.48	115.41

(b) Establishments

	2000-2010 data				2012 cross-section			
		— Average —		Median		— Average —		Median
# lyrs	est-yrs	# emp	wage	wage	est-yrs	# emp	wage	wage
0	166,358	8	83.08	80.88	11,647	6	74.43	68.89
1	88,206	21	106.59	103.80	14,585	11	95.68	88.58
2	32,041	65	108.74	104.62	7,433	71	108.17	104.71
3	14,360	466	119.46	119.03	1,415	268	115.86	114.82

This table is based on Table 3 in Caliendo et al. (2015). It reports descriptive statistics for HQ and establishments by their number of managerial layers. We separate the years 2000-2010 and 2012 because of different classifications of occupations (see Appendix A.2). *# of emp.* refers to the number of employees. *wage* refers to daily wages.

Table A.24: Data description by number of managerial layers in firm

(a) Headquarters

	2000-2010 data				2012 cross-section			
		— Average —		Median		— Average —		Median
# lyrs	HQ-yrs	# emp	wage	wage	HQ-yrs	# emp	wage	wage
0	11,939	25	79.12	77.97	957	15	70.61	68.83
1	20,112	36	91.00	89.14	1,865	22	82.00	77.98
2	20,683	71	101.28	100.18	3,313	71	98.32	96.29
3	24,840	307	116.22	116.01	3,677	207	113.04	113.14

(b) Establishments

	2000-2010 data				2012 cross-section			
		— Average —		Median		— Average —		Median
# lyrs	est-yrs	# emp	wage	wage	est-yrs	# emp	wage	wage
0	22,925	8	73.12	69.98	2,018	7	62.01	56.98
1	48,217	11	81.77	79.19	3,574	9	74.80	69.32
2	63,395	16	89.44	86.65	8,814	20	93.99	89.85
3	166,428	61	102.95	98.79	20,674	44	97.20	90.60

This table is based on Table 3 in Caliendo et al. (2015). It reports descriptive statistics for HQ and establishments by the number of managerial layers in the firm. We separate the years 2000-2010 and 2012 because of different classifications of occupations (see Appendix A.2). *# of emp.* refers to the number of employees. *wage* refers to daily wages.

Three findings are noteworthy. First, Table A.23 shows that HQ have a higher number of managerial layers than establishments. Second, units (i.e., establishments/HQ) with a higher number of layers are larger and pay higher average and median wages. Counting layers at the unit level or the firm level does not affect this pattern. Third, a comparison of Tables A.23 and A.24 shows that counting layers at the unit level or the firm level matters for establishment statistics, but less so for headquarter statistics. If the number of layers is determined at the unit level, establishments pay higher average wages than HQ with the same number of layers. If the number of layers is determined at the firm level, establishments pay lower wages than HQ with the same number of layers. The average size of establishments by layer also changes considerably.

To understand these patterns, it is useful to compare the number of layers in the firm and the establishments/HQ, as shown in Table A.25. Most establishments have fewer layers than the firm, while HQ typically have the same number of layers as the firm. Firms with more layers also tend to have more establishments (including small ones). Firms with more layers are thus overrepresented in the sample of establishments, but not in the sample of HQ. As firms with more layers tend to be larger and pay higher wages, establishment wages are considerably higher if layers are counted at the unit level than if layers are counted at the firm level, while headquarter wages are hardly affected.

Table A.25: Comparison: number of managerial layers in firm vs. HQ/establishment

(a) Headquarters

	2000-2010 data					2012 cross-section				
# of lyrs in firm	# of lyrs in HQ 0	1	2	3	# HQ-yrs	# of lyrs in HQ 0	1	2	3	# HQ
0	100				11,939	100				957
1	21	79			20,112	23	77			1,865
2	6	21	73		20,683	6	17	77		3,313
3	2	4	14	80	24,840	1	4	15	79	3,677

(b) Establishments

	2000-2010 data					2012 cross-section				
# of lyrs in firm	# of lyrs in est. 0	1	2	3	# est-yrs	# of lyrs in est. 0	1	2	3	# est.
0	100				22,925	100				2,018
1	70	30			48,217	58	42			3,574
2	51	36	13		63,395	32	40	28		8,814
3	47	31	14	9	166,428	23	46	24	7	20,674

This table reports the percentage share of HQ (establishments) with 0, ..., 3 managerial layers among firms with 0, ..., 3 managerial layers. We separate the years 2000-2010 and 2012 because of different classifications of occupations (see Appendix A.2).

Table A.26 reports the fraction of HQ and establishments with consecutively ordered layers for different numbers of managerial layers. Results are similar to results at the firm level reported in Table A.4. Unlike firms, not all establishments or HQ with three managerial layers have consecutive layers, because they may lack the layer of production workers (layer 0).

Table A.26: Percentage of HQ/establishments that have consecutively ordered layers

(a) Headquarters

	2000-2010 data Among HQ with					2012 cross-section Among HQ with				
	0 lyrs	1 lyr	2 lyrs	3 lyrs	All HQ	0 lyrs	1 lyr	2 lyrs	3 lyrs	All HQ
Unw.	99.88	54.49	28.39	99.95	70.44	97.49	59.69	76.39	99.18	82.99
W. by # emp.	99.94	58.55	37.55	100.00	86.51	98.55	63.65	90.96	99.99	95.97

(b) Establishments

	2000-2010 data Among est. with					2012 cross-section Among est. with				
	0 lyrs	1 lyr	2 lyrs	3 lyrs	All est.	0 lyrs	1 lyr	2 lyrs	3 lyrs	All est.
Unw.	99.97	50.36	44.41	99.53	79.49	99.18	30.93	71.47	96.68	64.83
W. by # emp.	99.98	65.29	59.34	99.99	87.50	99.20	51.05	93.39	99.68	90.07

This table is based on Table 4 in Caliendo et al. (2015). It reports the fraction of HQ/establishments with consecutively ordered layers for different numbers of managerial layers. The first row presents the unweighted share. The second row presents the shares weighted by the total number of employees. We separate the years 2000-2010 and 2012 because of different classifications of occupations (see Appendix A.2).

Tables A.27 to A.30 present the shares of HQ and establishments that satisfy a hierarchy in the number of employees and average wages respectively. Studying hierarchies at the level of the establishment or HQ is more complex than studying them at the firm level, because not all units have every layer of the firm. For example, in a firm with layers 0, 1, and 2, the establishment may have layer 0 and 1 and the headquarters only layer 0 and 2. Assessing whether the number of employees are higher and wages are lower at layer 0 than 1 is thus not straightforward for the HQ.

Hence, we construct two variants of the respective tables. Tables A.27 and A.29 count layers at the establishment or HQ level. We thus compare adjacent layers irrespective of their position within the firm. In the example, layer 2 in the HQ would thus be treated as layer 1. Tables A.28 and A.30 restrict the sample to those HQ and establishments that have the same sequence of layers as the firm. We would thus drop the HQ, but keep the establishment in the sample in our example. We keep 94% of HQ in the cross-section (93% in the panel), but only 64% of establishments (70% in the panel), reflecting that the organizational structure of HQ tends to be closer to that of the firm.

Table A.27: HQ/establishments that satisfy a hierarchy in the # employees

(a) Headquarters

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$
1	89.3	89.3			84.0	84.0		
2	70.4	89.2	81.2		58.7	76.5	78.8	
3	45.0	89.8	88.7	59.2	47.2	72.4	76.3	91.5

(b) Establishments

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$
1	82.4	82.4			83.7	83.7		
2	70.2	85.9	84.1		62.4	77.0	82.8	
3	42.0	86.6	88.6	68.5	46.8	76.3	73.9	90.3

This table is based on Table 5 in Caliendo et al. (2015). It reports the share of HQ/establishments that satisfy a hierarchy in the number of employees separately for HQ/establishments with 1, 2, and 3 managerial layers. Layers are counted at the HQ/establishment level. Column notes see Table A.5.

Table A.28: HQ/est. that satisfy a hierarchy in the # employees, restricted sample

(a) Headquarters

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$
1	88.6	88.6			83.4	83.4		
2	71.1	89.0	82.1		58.1	76.4	78.1	
3	45.0	89.8	88.7	59.2	47.2	72.4	76.3	91.5

(b) Establishments

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$	$n_L^\ell \geq n_L^{\ell+1}$ $\forall \ell$	$n_L^0 \geq n_L^1$	$n_L^1 \geq n_L^2$	$n_L^2 \geq n_L^3$
1	76.9	76.9			74.9	74.9		
2	70.7	88.3	82.2		59.7	74.8	82.1	
3	50.8	86.6	86.6	68.5	46.8	76.3	73.9	90.3

This table is based on Table 5 in Caliendo et al. (2015). It reports the share of HQ/establishments that satisfy a hierarchy in the number of employees separately for HQ/establishments with 1, 2, and 3 managerial layers. The sample includes only HQ/establishments with the same sequence of layers as the firm. Column notes see Table A.5.

Table A.29: HQ/establishments that satisfy a hierarchy in wages

(a) Headquarters

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	89.5	89.5			82.8	82.8		
2	67.2	91.0	75.7		66.6	83.3	82.1	
3	41.7	94.5	60.7	82.5	63.5	86.5	85.1	88.3

(b) Establishments

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	86.5	86.5			86.4	86.4		
2	66.1	87.9	77.5		65.6	82.6	81.6	
3	43.3	93.4	60.1	85.9	57.2	84.5	81.4	88.1

This table is based on Table 6 in Caliendo et al. (2015). It reports the share of HQ/establishments that satisfy a hierarchy in wages separately for HQ/establishments with 1, 2, and 3 managerial layers. Layers are counted at the HQ/establishment level. Column notes see Table A.6.

Table A.30: HQ/establishments that satisfy a hierarchy in wages, restricted sample

(a) Headquarters

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	89.3	89.3			81.9	81.9		
2	66.2	90.0	74.7		65.9	83.2	81.4	
3	41.7	94.5	60.7	82.5	63.5	86.5	85.1	88.3

(b) Establishments

	2000-2010 data				2012 cross-section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of layers	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^{\ell+1} \geq w_L^{\ell}$ $\forall \ell$	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	83.4	83.4			77.2	77.2		
2	59.3	87.2	71.3		64.9	82.1	81.4	
3	43.3	93.4	60.1	85.9	57.2	84.5	81.4	88.1

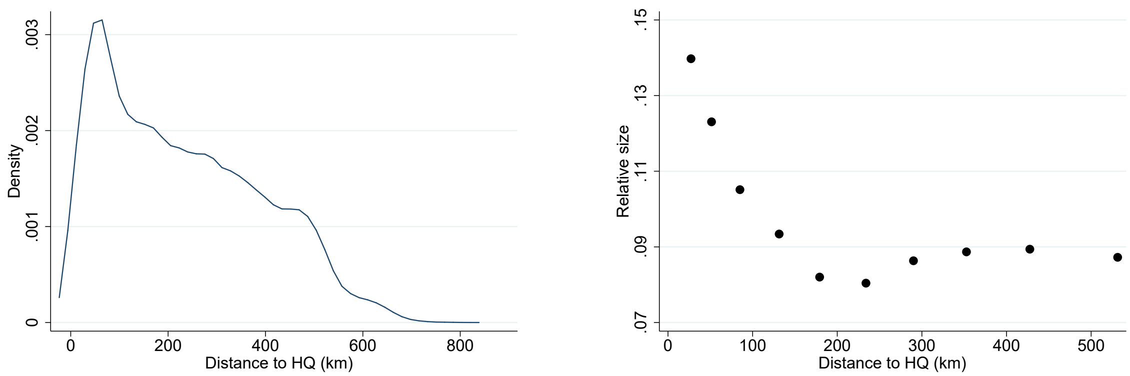
This table is based on Table 6 in Caliendo et al. (2015). It reports the share of headquarters/establishments that satisfy a hierarchy in wages separately for HQ/establishments with 1, 2, and 3 managerial layers. The sample includes only HQ/establishments with the same sequence of layers as the firm. Column notes see Table A.5.

B. Facts

B.1. Distance to headquarters decreases location probability

Figure B.1: Distance to HQ and location and size of establishments, 2012 cross section

(a) Location (b) Size



Left: density of establishments over distance to headquarters. Right: bin scatter plot of relation between relative size of establishments and the distance to headquarters, where relative size is defined as the number of employees of establishments relative to the number of employees of the firm. # establishments: 31,718.

Table B.1: Location probability and establishment size, ME firms, 2000-2010 data

Dependent variable	Location probability			Log # est. employees		
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	−0.316***	−0.314***	−0.379***	−0.109***	−0.107***	−0.146***
	(0.020)	(0.021)	(0.016)	(0.022)	(0.021)	(0.019)
Log market potential	0.744***	0.752***		0.521***	0.526***	
	(0.025)	(0.029)		(0.061)	(0.052)	
Relative wages	−1.120***	−0.857***		−0.385*	−0.363*	
	(0.057)	(0.059)		(0.172)	(0.159)	
Relative land prices		−0.016*			0.011	
		(0.006)			(0.007)	
# observations	24,394,362	11,824,458	24,394,362	171,146	86,084	171,146
# firms	10,323	8,478	10,323	8,547	6,982	8,547
HQ sector FE	Y	Y	Y	N	N	N
HQ county FE	Y	Y	Y	N	N	N
Legal form FE	Y	Y	Y	N	N	N
County FE	N	N	Y	N	N	Y
Year FE	Y	Y	Y	N	N	N
Firm-year FE	N	N	N	Y	Y	Y
Model		Probit			OLS	

The table presents the coefficient estimates of a probit model (standard errors clustered by HQ county in parentheses) in columns 1-3 and a linear model (standard errors clustered by firm and county in parentheses) in columns 4-6. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table III. Land prices are only available from 2005. Columns 4-6 only include firms establishments in at least two counties. FE = fixed effects.

B.2. Distance to headquarters increases the number of layers

Modifications to the main variables.

- Tables B.2 and B.3 replicate the analyses treating the lowest-level layer in each establishment as non-managerial.
- Tables B.4 and B.5 use dummies for the quartiles of the (maximum) distance to headquarters. We find that the effect of distance increases monotonically.
- Tables B.6 and B.7 include squared size, and Table B.8 includes the number of establishments as covariate. The results are robust, which suggests that distance does not merely take up omitted non-linear effects of size on the organization.
- In unreported regressions, we find that the results in Table V are robust to including the age of the establishment as covariate, so a higher prevalence of local managers in the set-up phase of an establishment does not drive the results.

Table B.2: Regression results, mg. organization of ME firms, establishment-level layer definition, 2012 cross-section

Dependent variable	# managerial layers				Mg. share $\in [0,1]$ Layers	
	(1)	(2)	(3)	(4)	(5)	(6)
Maximum log distance to HQ	0.037***		0.043***		0.130***	
	(0.007)		(0.006)		(0.014)	
Log area		0.025***		0.026***		0.068***
		(0.005)		(0.004)		(0.010)
Log sales	0.125***	0.093***				
	(0.004)	(0.006)				
Log # non-mg. employees			0.133***	0.113***		
			(0.004)	(0.006)		
# firms	4,323	1,661	7,742	2,768	7,742	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM	

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* (1)-(4) number of managerial layers, (5),(6) managerial share in wage sum, both defined treating the lowest-level layer in each establishment/the HQ as non-managerial. *Independent variables:* see Table IV.

Table B.3: Regression results, mg. organization of establishments, establishment-level layer definition, 2012 cross-section

Unit	Establishment		Headquarters	
Dependent variable	# layers	Mg. share $\in [0,1]$ Layers	# layers	Mg. share $\in [0,1]$ Layers
	(1)	(2)	(3)	(4)
Log distance to HQ	0.039^{***}	0.062^{**}		
	(0.011)	(0.021)		
Maximum log distance to HQ			0.058^{***}	0.173^{***}
			(0.007)	(0.015)
Log # non-mg. employees	0.309^{***}		0.174^{***}	
	(0.012)		(0.005)	
# est./HQ	31,717	31,717	8,217	8,217
Sector FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
Model	Poisson	GLM	Poisson	GLM

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 and 2, robust in columns 3 and 4). ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1),(3) number of managerial layers, (2),(4) managerial share in wage sum, both defined treating the lowest-level layer in each establishment/the HQ as non-managerial. *Independent variables:* see Table V.

Table B.4: Regression results, mg. organization of ME firms, distance quartiles, 2012 cross-section

Dependent variable	# managerial layers		Mg. share $\in [0,1]$ Layers	Mg. share $\in [0,1]$ Blossfeld
	(1)	(2)	(3)	(4)
Quartile 2	0.071^{***}	0.040^{*}	0.068^{+}	0.091^{+}
	(0.020)	(0.016)	(0.038)	(0.049)
Quartile 3	0.111^{***}	0.104^{***}	0.295^{***}	0.237^{***}
	(0.019)	(0.015)	(0.038)	(0.048)
Quartile 4	0.089^{***}	0.118^{***}	0.382^{***}	0.312^{***}
	(0.020)	(0.016)	(0.040)	(0.050)
Log sales	0.116^{***}			
	(0.004)			
Log # non-mg. employees		0.109^{***}		
		(0.004)		
# firms	4,323	7,742	7,742	7,742
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y
Model	Poisson		GLM	

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1),(2) number of managerial layers, (3) managerial share in wage sum, according to layers, (4) managerial share in wage sum, according to Blossfeld occupational categories. *Independent variables: Quartile 2-4*: dummies for quartiles of the log of maximum distance between establishment and headquarters in km; others see Table IV.

Table B.5: Regression results, mg. organization of establishments, distance quartiles, 2012 cross-section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance	0.115***	0.328***	0.242*			
quartile 2	(0.022)	(0.056)	(0.111)			
Log distance	0.132***	0.446***	0.457**			
quartile 3	(0.026)	(0.075)	(0.143)			
Log distance	0.127***	0.422***	0.531***			
quartile 4	(0.028)	(0.077)	(0.137)			
Max. log distance				0.027	0.047	0.097
quartile 2				(0.018)	(0.041)	(0.060)
Max. log distance				0.103***	0.275***	0.211***
quartile 3				(0.018)	(0.041)	(0.057)
Max. log distance				0.134***	0.414***	0.386***
quartile 4				(0.018)	(0.043)	(0.058)
Log # non-mg.	0.256***			0.162***		
employees	(0.010)			(0.004)		
# est./HQ	26,409	31,717	31,717	7,999	8,217	8,217
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* see Table V. *Independent variables: Log distance quartile 2-4*: dummies for quartiles of log of distance between establishment and headquarters in km; *Max. log distance quartile 2-4*: dummies for quartiles of log of maximum distance between establishment and headquarters in km; *Log # of non-mg. employees:* log number of employees at lowest layer.

Table B.6: Regression results, mg. organization of ME firms, non-linear size, 2012 cross-section

Dependent variable	# managerial layers			
	(1)	(2)	(3)	(4)
Maximum log distance to HQ	0.037***		0.049***	
	(0.007)		(0.006)	
Log area		0.019***		0.027***
		(0.004)		(0.004)
Log sales	0.478***	0.461***		
	(0.031)	(0.041)		
Log sales, squared	−0.017***	−0.017***		
	(0.001)	(0.002)		
Log # non-mg. employees			0.086***	0.109***
			(0.012)	(0.017)
Log # non-mg. employees, squared			0.003+	−0.002
			(0.001)	(0.002)
# firms	4,323	1,661	7,742	2,768
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y

The table presents the estimated coefficients of Poisson regressions. Constant included. Robust standard errors in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* number of managerial layers. *Independent variables:* see Table IV.

Table B.7: Regression results, mg. organization of establishments, non-linear size, 2012 cross-section

Unit	Est.	HQ
Dependent variable	# mg. layers	
	(1)	(2)
Log distance to HQ	0.056***	
	(0.011)	
Maximum log distance to HQ		0.059***
		(0.007)
Log # non-mg. employees	0.339***	0.217***
	(0.026)	(0.014)
Log # non-mg. employees, squared	−0.015***	−0.008***
	(0.004)	(0.002)
# est./HQ	26,409	7,999
Sector FE	Y	Y
County FE	Y	Y

The table presents the estimated coefficients of Poisson regressions. Constant included. Standard errors in parentheses (clustered by firm in column 1, robust in column 2). *** $p < 0.001$. *Dependent variable:* number of managerial layers. *Independent variables:* see Table V.

Table B.8: Regression results, mg. organization of ME firms, number of establishments, 2012 cross-section

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.039***		0.050***		0.152***		0.132***	
	(0.007)		(0.006)		(0.015)		(0.019)	
Log area		0.025***		0.027***		0.069***		0.072***
		(0.005)		(0.004)		(0.011)		(0.013)
Log sales	0.116***	0.083***						
	(0.004)	(0.005)						
Log # non-mg. employees			0.110***	0.088***				
			(0.004)	(0.005)				
# establishments	−0.000	−0.000	−0.000*	−0.000	0.000	0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# firms	4,323	1,661	7,742	2,768	7,742	2,768	7,742	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* see Table IV. *Independent variables: # establishments:* number of establishments (excluding HQ), others: see Table IV.

Alternative econometric specifications.

- Tables B.9 and B.10 replicate the regression results using linear models.
- Table B.11 shows that distance affects the managerial organization of an establishment even within firms: the number of layers and the managerial share of an establishment increase with distance in linear regressions including firm fixed effects.

Table B.9: Regression results, mg. organization of ME firms, OLS, 2012 cross-section

Dependent variable	# managerial layers				Mg. share (%)			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.069***		0.094***		2.789***		0.909***	
	(0.014)		(0.011)		(0.279)		(0.135)	
Log area		0.050***		0.054***		1.325***		0.555***
		(0.011)		(0.009)		(0.236)		(0.116)
Log sales	0.250***	0.192***						
	(0.008)	(0.013)						
Log # non-mg. employees			0.233***	0.200***				
			(0.008)	(0.013)				
# firms	4,270	1,529	7,713	2,673	7,713	2,673	7,713	2,673
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table IV.

Table B.10: Regression results, mg. organization of establishments, OLS, 2012 cross-section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share (%)		# layers	Mg. share (%)	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.052***	3.605***	1.414***			
	(0.010)	(0.646)	(0.412)			
Maximum log distance to HQ				0.102***	3.187***	1.308***
				(0.011)	(0.308)	(0.186)
Log # non-mg. employees	0.278***			0.303***		
	(0.009)			(0.008)		
# est./HQ	26,395	31,700	31,700	7,972	8,190	8,190
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y

Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). *** $p < 0.001$. *Dependent and independent variables:* see Table V.

Table B.11: Regression results, mg. organization of establishments within firms, OLS, 2012 cross-section

Dependent variable	# layers	Mg. share (%)		# layers	Mg. share (%)	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance	0.026***	1.163***	0.415***	0.025***	1.189***	0.468**
to HQ	(0.007)	(0.262)	(0.157)	(0.007)	(0.280)	(0.143)
Log # non-mg.	0.254***			0.260***		
employees	(0.010)			(0.009)		
# est./HQ	21,764	26,617	26,617	21,784	26,636	26,636
Sector FE	Y	Y	Y	N	N	N
County FE	Y	Y	Y	N	N	N
Firm FE	Y	Y	Y	Y	Y	Y

Standard errors clustered by firm in parentheses. ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table V.

2000–2010 panel and sample splits.

- Tables B.12 and B.13 show that the results are similar in the 2000–2010 data.
- Tables B.14 and B.15 show that the results are similar for firms that found establishments for horizontal and vertical motives. We approximate motives using the sector: establishments in the same sector as the headquarters are considered horizontal; establishments in a different sector are considered vertical. In unreported regressions, we use the main non-managerial occupation as the criterion and obtain similar results.
- Tables B.16 and B.17 present the results by the legal form of the firm. The legal form affects whether owner-managers have to contribute to social security and are thus included in the data. Establishment-level results are robust for all legal forms. Firm-level and headquarter-level results are robust, except for public companies. This is unsurprising, given that there are very few public companies in the sample and more than 90 percent of these firms and almost 90 percent of their headquarters have two or three layers, hence there is little variation in their managerial organization.

Table B.12: Regression results, mg. organization of ME firms, 2000-2010 data

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.068^{***}		0.071^{***}		0.146^{***}		0.075^{**}	
	(0.008)		(0.007)		(0.016)		(0.024)	
Log area		0.032^{***}		0.038^{***}		0.058^{***}		0.047^{*}
		(0.005)		(0.005)		(0.013)		(0.018)
Log sales	0.140^{***}	0.115^{***}						
	(0.004)	(0.007)						
Log # non-mg. employees			0.183^{***}	0.152^{***}				
			(0.005)	(0.007)				
# firm-years	18,782	7,383	33,882	12,257	33,882	12,257	33,882	12,257
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficients. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with at least two establishments (plus headquarters). *Dependent and independent variables:* see Table IV.

Table B.13: Regression results, mg. organization of establishments, 2000-2010 data

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.073^{***}	0.134^{***}	0.068^{+}			
	(0.011)	(0.037)	(0.037)			
Maximum log distance to HQ				0.076^{***}	0.144^{***}	0.128^{***}
				(0.007)	(0.015)	(0.023)
Log # non-mg. employees	0.353^{***}			0.242^{***}		
	(0.019)			(0.005)		
# est./HQ-years	259,723	274,433	274,433	64,269	64,691	64,691
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

Standard errors clustered by firm in parentheses. $^{+}$ $p < 0.1$, *** $p < 0.001$. *Dependent and independent variables:* see Table V.

Table B.14: Regression results, mg. organization of ME firms, by investment motive, 2012 cross-section

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
Horizontal	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.038***		0.046***		0.136***		0.124***	
	(0.011)		(0.008)		(0.021)		(0.025)	
Log area		0.019*		0.027***		0.071***		0.060**
		(0.008)		(0.007)		(0.020)		(0.023)
Log sales	0.124***	0.105***						
	(0.007)	(0.011)						
Log # non-mg. employees			0.113***	0.100***				
			(0.006)	(0.012)				
# firms	2,190	716	3,987	1,249	3,987	1,249	3,987	1,249
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

Vertical	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.031**		0.049***		0.167***		0.137***	
	(0.011)		(0.009)		(0.022)		(0.028)	
Log area		0.022***		0.021***		0.057***		0.075***
		(0.006)		(0.006)		(0.016)		(0.017)
Log sales	0.110***	0.070***						
	(0.005)	(0.006)						
Log # non-mg. employees			0.105***	0.076***				
			(0.005)	(0.006)				
# firms	2,133	945	3,755	1,519	3,755	1,519	3,755	1,519
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table IV. *Horizontal* restricts the sample to firms with all establishments in the same sector as the HQ. *Vertical* restricts the sample to firms with at least one establishment in a different sector than the HQ.

Table B.15: Regression results, mg. organization of establishments, by investment motive, 2012 cross-section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
Horizontal	(1)	(2)	(3)	(4)	(5)	(6)
Log distance	0.075***	0.232***	0.281***			
to HQ	(0.017)	(0.046)	(0.050)			
Maximum log				0.051***	0.165***	0.158***
distance to HQ				(0.009)	(0.022)	(0.030)
Log # non-mg.	0.262***			0.154***		
employees	(0.013)			(0.007)		
# est./HQ	9,365	10,873	10,873	4,120	4,229	4,229
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

Vertical	(1)	(2)	(3)	(4)	(5)	(6)
Log distance	0.041**	0.149***	0.211**			
to HQ	(0.013)	(0.039)	(0.060)			
Maximum log				0.063***	0.176***	0.167**
distance to HQ				(0.010)	(0.024)	(0.034)
Log # non-mg.	0.254***			0.167***		
employees	(0.013)			(0.006)		
# est./HQ	17.044	20,844	20,844	3,879	3,988	3,988
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table V. *Horizontal* restricts the sample to establishments and HQ of firms with all establishments in the HQ sector. *Vertical* restricts the sample to establishments and HQ of firms with at least one establishment in a different sector than the HQ.

Table B.16: Regression results, mg. organization of ME firms, by legal form, 2012 cross-section

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
GmbH & Co. KG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.032		0.019		0.095**		0.132**	
	(0.022)		(0.014)		(0.034)		(0.043)	
Log area		0.018		0.049***		0.151***		0.090*
		(0.017)		(0.013)		(0.034)		(0.036)
Log sales	0.148***	0.073***						
	(0.014)	(0.020)						
Log # non-mg. employees			0.144***	0.085***				
			(0.011)	(0.019)				
# firms	589	215	1,209	452	1,209	452	1,209	452
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

GmbH	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.046***		0.059***		0.165***		0.136***	
	(0.009)		(0.007)		(0.017)		(0.022)	
Log area		0.031***		0.029***		0.070***		0.064***
		(0.006)		(0.005)		(0.014)		(0.016)
Log sales	0.125***	0.105***						
	(0.005)	(0.007)						
Log # non-mg. employees			0.110***	0.090***				
			(0.005)	(0.007)				
# firms	3,360	1,212	6,011	2,018	6,011	2,018	6,011	2,018
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

AG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	−0.007		−0.003		0.114		0.155+	
	(0.025)		(0.020)		(0.078)		(0.092)	
Log area		−0.004		−0.003		−0.054		0.014
		(0.011)		(0.010)		(0.044)		(0.056)
Log sales	0.021*	0.015						
	(0.009)	(0.011)						
Log # non-mg. employees			0.046***	0.041***				
			(0.008)	(0.010)				
# firms	363	228	506	291	506	291	506	291
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The tables present the coefficient estimates separately for firms with the legal form *GmbH*, *GmbH & Co. KG* and *AG*. Constant included. Robust standard errors in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table IV. A "GmbH" is a limited liability company. A "GmbH & Co. KG" is a limited partnership with a limited liability company as general partner. "AGs" are public companies.

Table B.17: Regression results, mg. organization of establishments, by legal form, 2012 cross-section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
GmbH & Co. KG	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.087**	0.106	0.318**			
	(0.023)	(0.065)	(0.066)			
Maximum log distance to HQ				0.048**	0.129**	0.150**
				(0.017)	(0.040)	(0.055)
Log # non-mg. employees	0.336***			0.196***		
	(0.027)			(0.014)		
# est./HQ	4,233	4,704	4,704	1,191	1,209	1,209
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

GmbH	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.063***	0.203***	0.272*			
	(0.014)	(0.043)	(0.072)			
Maximum log distance to HQ				0.067***	0.163***	0.164***
				(0.009)	(0.019)	(0.027)
Log # non-mg. employees	0.258***			0.173***		
	(0.011)			(0.006)		
# est./HQ	15,727	18,343	18,343	5,840	6,011	6,011
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

AG	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.045*	0.197**	0.132*			
	(0.019)	(0.060)	(0.058)			
Maximum log distance to HQ				0.038	0.127	0.148
				(0.028)	(0.091)	(0.106)
Log # non-mg. employees	0.288***			0.065***		
	(0.024)			(0.013)		
# est./HQ	3,679	5,586	5,586	492	506	506
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y

The tables present the coefficient estimates separately for establishments/HQ of firms with the legal form *GmbH*, *GmbH & Co. KG* and *AG*. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table IV. A "GmbH" is a limited liability company. A "GmbH & Co. KG" is a limited partnership with a limited liability company as general partner. "AGs" are public companies.

B.3. Reorganization of headquarters or establishments

Overview of Appendix tables.

- Table B.18 documents that changes in the number of layers for SE firms are similar to changes documented for French and Danish firms (Caliendo et al. 2015, Friedrich 2020).
- Table B.19 shows that changes in the number of layers are related to changes in firm size, consistent with the literature (Caliendo et al. 2015, Friedrich 2020).
- Table B.20 documents that the organizational dynamics are similar if the lowest-level layer in each establishment is treated as non-managerial.
- Table B.21 shows that the transition dynamics are comparable for longer time lags. Firms also typically add or drop layers either at the establishments or the headquarters when we consider a five-year period.
- Table B.22 shows that the results are similar for firms with headquarters and exactly one establishment, and for firms with headquarters and at least two establishments, hence the aggregation of the establishments in Table VI does not drive the results.
- Table B.23 documents that the transition patterns are similar for firms with only proximate and firms with distant establishments.

Table B.18: Transition dynamics of the managerial organization, SE firms, 2000-2010 data

# layers in $t/t+1$	0	1	2	3	ME	Exit	# firms
0	**92**	7				1	159,058
1	5	**87**	7		1		195,573
2		9	**83**	6	1		127,793
3		1	10	**88**	1		73,165

The table displays the percentage share of SE firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers in year $t+1$ (given in the columns). Empty cells contain fewer than .5% of observations. Diagonal in bold.

Table B.19: Size at transition, ME firms, 2000-2010 data

(a) firm

# layers in $t/t+1$	0	1	2	3	SE	# firms
0	**3.4**	3.5***	3.8***		3.3***	10,778
1	3.6***	**3.7**	3.9***	4.2	3.5***	18,274
2		4.0***	**4.3**	4.5***	3.9***	18,754
3			4.6***	**5.5**	4.8***	22,391

(b) headquarters/establishments

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**3.4**	3.6***						3.3***	10,778
HQ 1/ est. 0	3.7	**3.7**	3.8*	3.9***				3.5***	8,340
HQ 1/ est. 1	..	3.9	**3.8**	4.1***		..		3.6***	8,052
HQ 2/ est. 0,1		4.0***	4.1***	**4.3**	4.5***	4.4***		4.0***	12,046
HQ 2/ est. 2			..	4.6**	**4.8**	4.8	..	4.2***	3,410
HQ 3/ est. 0,1,2				4.5***	5.0***	**5.2**	5.8***	4.8***	13,365
HQ 3/ est. 3						6.0***	**6.7**	..	4,625

Panel (a) displays the average log number of employees of firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers in year $t+1$ (given in the columns). Panel (b) displays the average log number of employees of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the headquarters. The figure behind the slash denotes the maximum number of layers at the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. The stars denote whether average size of firms that change their organization is significantly different from the average size of those that do not (marked in bold). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. .. denotes cells with fewer than 50 observations. Empty cells contain fewer than .5% of firms. Fewer than .5% of firms exit. Unreported results with sales as outcome variable are similar.

Table B.20: Transition dynamics of the managerial organization, establishment-level layer definition, 2000-2010 data

(a) # managerial layers of firm

# layers in $t/t+1$	0	1	2	3	SE	# firms
0	**85**	8	1		6	10,968
1	5	**82**	7		6	20,327
2		8	**79**	7	5	18,696
3			6	**90**	4	20,206

(b) # managerial layers at headquarters/establishment

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**85**	5						6	10,968
HQ 1/ est. 0	6	**74**	4	6				8	9,252
HQ 1/ est. 1	1	6	**75**	7		1		3	7,006
HQ 2/ est. 0,1		4	4	**76**	2	6		7	12,144
HQ 2/ est. 2			1	10	**69**	9	1	2	3,254
HQ 3/ est. 0,1,2				5	2	**84**	3	5	13,374
HQ 3/ est. 3						9	**86**	1	4,606

Panel (a) displays the percentage share of firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers in year $t+1$ (given in the columns). Panel (b) displays the percentage share of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the headquarters. The figure behind the slash denotes the maximum number of layers at the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. Empty cells contain fewer than .5% of firms. Fewer than .5% of firms exit. Diagonal in bold.

Table B.21: Transition dynamics of the managerial organization, five-year time period, 2000-2010 data

(a) # managerial layers of firm

# layers in $t/t+5$	0	1	2	3	SE	# firms
0	**59**	16	3	1	21	6,272
1	10	**51**	14	3	21	10,954
2		14	**48**	16	19	11,211
3		2	11	**71**	15	13,044

(b) # managerial layers at headquarters/establishment(s)

# layers in $t/t+5$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**59**	10	2	2				21	6,272
HQ 1/ est. 0	12	**40**	7	10				26	4,977
HQ 1/ est. 1	2	8	**45**	11		3		16	4,821
HQ 2/ est. 0,1	1	7	7	**43**	4	12		22	7,204
HQ 2/ est. 2			3	16	**32**	18	4	10	2,014
HQ 3/ est. 0,1,2		1	1	10	3	**59**	5	18	7,768
HQ 3/ est. 3				1	2	20	**64**	5	2,757

Panel (a) displays the percentage share of firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers in year $t+5$ (given in the columns). Panel (b) displays the percentage share of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+5$ (given in the columns). The figure in front of the slash denotes the number of layers of the HQ. The figure behind the slash denotes the maximum number of layers of the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. Empty cells contain fewer than .5% of firms. Fewer than 1% of firms exit. Diagonal in bold. Among all firms that reorganize, 42% change the number of layers only at the headquarters, 39% change it only at the establishment, and 20% change it at both.

Table B.22: Transition dynamics of the managerial organization, by number of establishments, 2000-2010 panel

(a) ME firms with headquarters and one establishment

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**83**	5						8	7,945
HQ 1/ est. 0	6	**74**	4	6				10	6,339
HQ 1/ est. 1		4	**76**	7		1		6	5,330
HQ 2/ est. 0,1		4	4	**74**	2	6		9	8,388
HQ 2/ est. 2			.	10	**68**	8	.	3	1,665
HQ 3/ est. 0,1,2				5	2	**83**	2	8	8,276
HQ 3/ est. 3					.	11	**82**	2	1,410

(b) ME firms with headquarters and at least two establishments

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**90**	5						1	2,833
HQ 1/ est. 0	8	**78**	5	6		.		2	2,001
HQ 1/ est. 1	.	5	**77**	7	.	.		1	2,722
HQ 2/ est. 0,1		3	5	**79**	4	6		1	3,658
HQ 2/ est. 2			.	9	**70**	10	.		1,745
HQ 3/ est. 0,1,2				4	3	**87**	4	1	5,089
HQ 3/ est. 3						8	**88**		3,215

The table displays the percentage share of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the HQ. The figure behind the slash denotes the maximum number of layers at the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. Empty cells contain fewer than .5% of firms. Dots mark cells that contain more than .5%, but fewer than 20 observations, so are omitted for confidentiality. Fewer than .5% of firms exit. Diagonal in bold. Panel (a) contains firms that maintain HQ and exactly one establishment in year t. Panel (b) contains firms that maintain HQ and at least two establishments in year t.

Table B.23: Transition dynamics of the managerial organization, by median maximum establishment distance, 2000-2010 data

(a) ME firms with maximum establishment distance of up to 170 km

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**85**	5						7	7,054
HQ 1/ est. 0	6	**75**	4	5				8	5,226
HQ 1/ est. 1	.	5	**76**	6	1			4	4,208
HQ 2/ est. 0,1		5	4	**75**	2	5		8	5,985
HQ 2/ est. 2			.	11	**67**	8		3	1,449
HQ 3/ est. 0,1,2				5	2	**83**	2	7	4,945
HQ 3/ est. 3						12	**82**	.	1,122

(b) ME firms with maximum establishment distance above 170 km

# layers in $t/t+1$	0/0	1/0	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ est. 0	**85**	5	1					5	3,724
HQ 1/ est. 0	7	**73**	5	7		.		7	3,844
HQ 1/ est. 1	.	4	**77**	7		1		3	3,114
HQ 2/ est. 0,1		3	4	**76**	3	7		6	6,061
HQ 2/ est. 2			.	9	**70**	10	.	1	1,961
HQ 3/ est. 0,1,2				5	2	**85**	3	4	8,420
HQ 3/ est. 3						8	**88**	1	3,503

The table displays the percentage share of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the HQ. The figure behind the slash denotes the maximum number of layers at the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. Empty cells contain fewer than .5% of firms. Dots mark cells that contain more than .5%, but fewer than 20 observations, so are omitted for confidentiality. Fewer than .5% of firms exit. Diagonal in bold. We split the sample at the median of the maximum log distance of establishments from the headquarters (170 km). Panel (a) contains firms with all establishments within a distance of 170 km in year t. Panel (b) contains firms with establishments above the distance of 170 km in year t.

C. Model

C.1. Set-up

C.1.1. Assumption 1 and its implications

Assumption 1. *The predictability of the production process* λ*, the helping costs* θ_{00} *and the learning costs* c *are such that*

$$\lambda\theta_{00} > c \quad \text{and} \quad \lambda > 2c.$$

Assumption 1 is sufficient to ensure that knowledge levels are unequal across layers. Before we formally state and prove this in Proposition C.1, we prove the following Lemma:

Lemma 1. *Given the number of layers* $L \in \{0, 1, 2\}$*, under the parameter restriction in Assumption 1, as the multiplier on the output constraint* $\xi_{0,L}$ *decreases, the constraint* $z^{\ell}_{0,L} \geq z^{\ell-1}_{0,L}$ *first binds at layer* L *and then at layer* $L-1$*; then the constraint* $z^0_{0,L} \geq 0$ *binds and finally the constraint* $\bar{z}_{0,L} \geq z^L_{0,L}$ *binds.*

Proof. The multiplier $\xi_{0,L}$ is higher if the constraint $z^0_{0,L} \geq 0$ binds than if the constraint $\bar{z}_{0,L} \geq z^L_{0,L}$ binds $\forall L \geq 0$:

$$\begin{aligned}
\xi_{0,L} &= \frac{w_0 c\theta_{00}}{\lambda} e^{\lambda(\bar{z}_{0,L} - z^L_{0,L})} \quad \text{(from equations C.2.2, C.2.8)}\\
\xi_{0,L} &= \frac{w_0 c\theta_{00}}{\lambda} \quad \text{at } \bar{z}_{0,L} = z^L_{0,L}\\
\xi_{0,L} &= \frac{w_0\left(1 + cz^0_{0,L} + \frac{c}{\lambda} + \mathbb{1}(L \geq 1)\frac{c}{\lambda}\theta_{00}\sum_{\ell=1}^{L} e^{-\lambda z^{\ell-1}_{0,L}}\right)}{1 - e^{-\lambda \bar{z}_{0,L}}} \quad \text{(from equation C.2.6)}\\
\xi_{0,L} &\geq \frac{w_0\left(1 + \frac{c}{\lambda} + \mathbb{1}(L \geq 1)\frac{c}{\lambda}\theta_{00}\right)}{1 - e^{-\lambda \bar{z}_{0,L}}} > \frac{w_0 c\theta_{00}}{\lambda} \quad \text{at } z^0_{0,L} = 0 \quad \text{by } \theta_{00} < 1, \frac{1}{1 - e^{-\lambda \bar{z}_{0,L}}} > 1
\end{aligned}$$

For $L > 0$, $d\xi_{0,L}/dz^0_{0,L} > 0$. The constraint $z^1_{0,L} \geq z^0_{0,L}$ binds at $z^0_{0,L} > 0$ and thus at a value of $\xi_{0,L}$ that is higher than its value at $z^0_{0,L} = 0$:

$$\begin{aligned}
z^1_{0,L} =& \frac{1}{\lambda\theta_{00}} e^{\lambda z^0_{0,L}} - \frac{1}{c} \geq z^0_{0,L} \quad \Leftrightarrow \quad \frac{1}{\lambda\theta_{00}} e^{\lambda z^0_{0,L}} - \frac{1}{c} - z^0_{0,L} \geq 0\\
\text{At } z^0_{0,L} = 0: \quad & \frac{1}{\lambda\theta_{00}} - \frac{1}{c} < 0 \quad \text{by Assumption 1}\\
\frac{d}{dz^0_{0,L}} =& \frac{1}{\theta_{00}} e^{\lambda z^0_{0,L}} - 1 > 0 \, \forall z^0_{0,L} \geq 0 \text{ by } \theta_{00} < 1\\
\Rightarrow \quad z^1_{0,L} =& \frac{1}{\lambda\theta_{00}} e^{\lambda z^0_{0,L}} - \frac{1}{c} = z^0_{0,L} \text{ at } z^0_{0,L} > 0
\end{aligned}$$

The constraint $z^2_{0,L} \geq z^1_{0,L}$ binds at $z^1_{0,L} > z^0_{0,L}$ and thus an even higher value of $\xi_{0,L}$:

$$\begin{aligned}
z^2_{0,2} =& \frac{1}{\lambda} e^{\lambda(z^1_{0,2} - z^0_{0,2})} - \frac{1}{c} \geq z^1_{0,2} \quad \Leftrightarrow \quad \frac{1}{\lambda} e^{\lambda(z^1_{0,2} - z^0_{0,2})} - \frac{1}{c} - z^1_{0,2} \geq 0\\
\text{At } z^1_{0,2} = z^0_{0,2}: \quad & \frac{1}{\lambda} - \frac{1}{c} - z^1_{0,2} < 0 \quad \text{by Assumption 1}\\
\frac{d}{dz^1_{0,2}} =& e^{\lambda(z^1_{0,2} - z^0_{0,2})}\left(1 - \frac{dz^0_{0,2}}{dz^1_{0,2}}\right) - 1 > 0 \text{ by Assumption 1 } (\lambda > 2c)
\end{aligned}$$

$$\Rightarrow \quad \frac{1}{\lambda} e^{\lambda(z_{0,2}^{1} - z_{0,2}^{0})} - \frac{1}{c} = z_{0,2}^{1} \text{ at } z_{0,2}^{1} > z_{0,L}^{0}$$

□

Proposition C.1. *Given Assumption 1, the knowledge of agents at layer $\ell = 1, ..., L$ exceeds the knowledge of agents at layer $\ell - 1$ in an organization with $L \in \{1, 2\}$ managerial layers, i.e., the constraint $z_{0,L}^{\ell} \geq z_{0,L}^{\ell-1}$ never binds.*

Proof. Consider an organization with $L > 0$ managerial layers. Assume that the organization adds a layer of middle managers L' with $z_{0,L'}^{L'} = z_{0,L'}^{L}$. The new organization is equivalent to the organization with L layers, but one extra manager at layer L. This implies that the constraint $n_{0,L}^{L} \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{L-1}}$ is slack, so the organization is not optimal. The constraint $z_{0,L}^{\ell} \geq z_{0,L}^{\ell-1}$ therefore never binds at layer L. Lemma 1 implies that it never binds at layer $\ell < L$. □

C.1.2. Overlapping vs. non-overlapping knowledge

We assume that knowledge levels are overlapping, because non-overlapping knowledge levels render the analysis of multi-establishment firm organization considerably more cumbersome. We expect that our main results hold in a model with non-overlapping knowledge, although some details may differ.

In a single-establishment firm, neither overlaps nor gaps between the knowledge of two adjacent layers occur (only by assumption, see Garicano, 2000). Overlaps are not optimal because the overlapping portion of knowledge is never used at the higher layer. Erasing the overlap saves costs at constant output. Gaps are not optimal due to the downward sloping problem density. Shifting the knowledge of the higher layer downwards until the gap is closed increases output at constant costs.

In a multi-establishment firm, these insights apply to the below-CEO layers within the headquarters or the establishment. The arguments also imply that gaps or overlaps between below-CEO and CEO knowledge at *both* units cannot be optimal. It is possible, however, that there are overlaps between below-CEO and CEO knowledge at the headquarters or the establishment, because the CEO uses the overlapping knowledge to solve problems at the other unit. Likewise, it is not possible to exclude the possibility of gaps between below-CEO and CEO knowledge at either the headquarters or the establishment. Closing the gap would increase output at this unit at constant costs, but decrease output at the other unit, because it would reduce the total amount of knowledge available to that unit.

To solve the model, one would have to keep track of possible overlaps or gaps at a unit. One would also have to employ a more complex function to compute output per unit of labor input:

$$q_{j,\omega} = \begin{cases} n_{j,\omega}^{0} \left(1 - e^{-\lambda \bar{Z}_{0,\omega}}\right) & \text{if } Z_{j,\omega}^{L_j} \geq \underline{Z}_{0,\omega} \\ n_{j,\omega}^{0} \left(1 - e^{-\lambda Z_{j,\omega}^{L_j}} + e^{-\lambda \underline{Z}_{0,\omega}} - e^{-\lambda \bar{Z}_{0,\omega}}\right) & \text{if } Z_{j,\omega}^{L_j} \leq \underline{Z}_{0,\omega} \end{cases}$$

where $\underline{Z}_{0,\omega}, \bar{Z}_{0,\omega}$ denote the lower and upper ends of the CEO knowledge interval and $Z_{j,\omega}^{L_j}$ denotes the upper end of below-CEO knowledge at a unit.

C.2. Single-establishment firm organization

C.2.1. Results for $L \geq 2$ layers of middle managers

Constraints (5) and (6) apply to layer $\ell = 2, ..., L$. Managerial knowledge at layers $\ell = 1, ..., L-1$ is a recursive function of the knowledge level at the highest below-CEO layer:

$$e^{\lambda\left(z_{0,L}^{\ell-1} - z_{0,L}^{\ell-2}\right)} = \left(1 + cz_{0,L}^{\ell}\right)\frac{\lambda}{c} \quad \forall \ell = 2, ..., L. \tag{C.2.1}$$

The marginal production costs and the marginal benefit of CEO time are given by:

$$\xi_{0,L} = \frac{w_0\left(1 + cz_{0,L}^0 + \frac{c}{\lambda} + \mathbb{1}(L \geq 2)\theta_{00}\frac{c}{\lambda}\sum_{\ell=1}^{L} e^{-\lambda z_{0,L}^{\ell-1}}\right)}{1 - e^{-\lambda \bar{z}_{0,L}}},$$

$$\varphi_{0,L} = \frac{w_0 c}{\lambda} e^{\lambda\left(z_{0,L}^L - z_{0,L}^{L-1}\right)}.$$

For the comparative statics, we restrict our attention to firms with $L \leq 2$ layers of middle managers in line with sections II and III.

Proposition C.2. *Proposition 1 applies to a firm with $L = 2$ layers of middle managers.*

Proof. See Appendix C.2.3. □

C.2.2. Lagrangian equation and first order conditions

We use equation (5), which is binding in optimum, to substitute for $n_{0,L}^{\ell}$, $L \geq \ell > 0$.

$$\begin{aligned}
\mathcal{L} = {} & n_{0,L}^0 w_0\left(1 + cz_{0,L}^0\right) + \mathbb{1}(L \geq 1) n_{0,L}^0 \sum_{\ell=1}^{L} \theta_{00} e^{-\lambda z_{0,L}^{\ell-1}} w_0\left(1 + cz_{0,L}^{\ell}\right) + w_0\left(1 + c\bar{z}_{0,L}\right) \\
& + \xi_{0,L}\left(\tilde{q} - n_{0,L}^0\left(1 - e^{-\lambda \bar{z}_{0,L}}\right)\right) + \varphi_{0,L}\left(n_{0,L}^0 \theta_{00} e^{-\lambda z_{0,L}^L} - 1\right) \\
& + \bar{\eta}_{0,L}^{L+1}\left(z_{0,L}^L - \bar{z}_{0,L}\right) + \mathbb{1}(L \geq 1)\sum_{\ell=1}^{L} \bar{\eta}_{0,L}^{\ell}\left(z_{0,L}^{\ell-1} - z_{0,L}^{\ell}\right) - \bar{\eta}_{0,L}^0 z_{0,L}^0 - \eta_{0,L}^0 n_{0,L}^0
\end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial \bar{z}_{0,L}} = w_0 c - \xi_{0,L} n_{0,L}^0 \lambda e^{-\lambda \bar{z}_{0,L}} - \bar{\eta}_{0,L}^{L+1} = 0 \tag{C.2.2}$$

$$\frac{\partial \mathcal{L}}{\partial z_{0,L}^L} \begin{cases} \overset{L=0}{=} n_{0,0}^0\left(w_0 c - \varphi_{0,0}\theta_{00}\lambda e^{-\lambda z_{0,0}^0}\right) + \bar{\eta}_{0,0}^1 - \bar{\eta}_{0,0}^0 = 0 \\ \overset{L\geq 1}{=} n_{0,L}^0\left(w_0 c\theta_{00} e^{-\lambda z_{0,L}^{L-1}} - \varphi_{0,L}\theta_{00}\lambda e^{-\lambda z_{0,L}^L}\right) + \bar{\eta}_{0,L}^{L+1} - \bar{\eta}_{0,L}^L = 0 \end{cases} \tag{C.2.3}$$

$$\frac{\partial \mathcal{L}}{\partial z_{0,L}^{\ell}} = n_{0,L}^0 w_0\left(c\theta_{00} e^{-\lambda z_{0,L}^{\ell-1}} - \lambda\theta_{00} e^{-\lambda z_{0,L}^{\ell}}\left(1 + cz_{0,L}^{\ell+1}\right)\right) - \bar{\eta}_{0,L}^{\ell} + \bar{\eta}_{0,L}^{\ell+1} = 0 \tag{C.2.4}$$

for $L > \ell > 0$, $L \geq 2$

$$\frac{\partial \mathcal{L}}{\partial z_{0,L}^0} \overset{L\geq 1}{=} n_{0,L}^0 w_0\left(c - \lambda\theta_{00} e^{-\lambda z_{0,L}^0}\left(1 + cz_{0,L}^1\right)\right) + \bar{\eta}_{0,L}^1 - \bar{\eta}_{0,L}^0 = 0 \tag{C.2.5}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial n_{0,L}^0} = {} & w_0\left(1 + cz_{0,L}^0 + \mathbb{1}(L \geq 1)\sum_{\ell=1}^{L} \theta_{00} e^{-\lambda z_{0,L}^{\ell-1}}\left(1 + cz_{0,L}^{\ell}\right)\right) \\
& - \xi_{0,L}\left(1 - e^{-\lambda \bar{z}_{0,L}}\right) + \varphi_{0,L}\theta_{00} e^{-\lambda z_{0,L}^L} - \eta_{0,L}^0 = 0
\end{aligned} \tag{C.2.6}$$

$$\text{(C.2.7)} \quad \frac{\partial \mathcal{L}}{\partial \xi_{0,L}} = \tilde{q} - n^0_{0,L}\left(1 - e^{-\lambda \bar{z}_{0,L}}\right) = 0$$

$$\text{(C.2.8)} \quad \frac{\partial \mathcal{L}}{\partial \varphi_{0,L}} = n^0_{0,L}\theta_{00}e^{-\lambda z^L_{0,L}} - 1 = 0$$

C.2.3. Propositions 1, C.2: Comparative statics

We differentiate the first order conditions with respect to output and solve the resulting system of linear equations. The second order conditions are:

$$\text{(C.2.9)} \quad \frac{d^2\mathcal{L}}{d\bar{z}_{0,L}d\tilde{q}} = -\frac{d\xi_{0,L}}{d\tilde{q}}n^0_{0,L}e^{-\lambda\bar{z}_{0,L}} - \xi_{0,L}\frac{dn^0_{0,L}}{d\tilde{q}}e^{-\lambda\bar{z}_{0,L}} + \xi_{0,L}n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}} = 0$$

$$\text{(C.2.10)} \quad \frac{d^2\mathcal{L}}{dz^L_{0,L}d\tilde{q}} \begin{cases} \overset{L=0}{=} -\frac{d\varphi_{0,0}}{d\tilde{q}}\theta_{00}\lambda e^{-\lambda z^0_{0,0}} + \varphi_{0,0}\theta_{00}\lambda^2 e^{-\lambda z^0_{0,0}}\frac{dz^0_{0,0}}{d\tilde{q}} = 0 \\ \overset{L\geq 1}{=} -w_0 c\lambda e^{-\lambda z^{L-1}_{0,L}}\frac{dz^{L-1}_{0,L}}{d\tilde{q}} - \frac{d\varphi_{0,L}}{d\tilde{q}}\lambda e^{-\lambda z^L_{0,L}} + \varphi_{0,L}\lambda^2 e^{-\lambda z^L_{0,L}}\frac{dz^L_{0,L}}{d\tilde{q}} = 0 \end{cases}$$

$$\frac{d^2\mathcal{L}}{dz^\ell_{0,L}d\tilde{q}} = -\lambda c e^{-\lambda z^{\ell-1}_{0,L}}\frac{dz^{\ell-1}_{0,L}}{d\tilde{q}} + \lambda^2 e^{-\lambda z^\ell_{0,L}}\frac{dz^\ell_{0,L}}{d\tilde{q}}(1 + cz^{\ell+1}_{0,L}) - \lambda e^{-\lambda z^\ell_{0,L}}c\frac{dz^{\ell+1}_{0,L}}{d\tilde{q}} = 0$$

$$\text{(C.2.11)} \quad \text{for } L > \ell > 0,\ L \geq 2$$

$$\text{(C.2.12)} \quad \frac{d^2\mathcal{L}}{dz^0_{0,L}d\tilde{q}} \overset{L\geq 1}{=} \lambda^2\theta_{00}e^{-\lambda z^0_{0,L}}\frac{dz^0_{0,L}}{d\tilde{q}}(1 + cz^1_{0,L}) - \lambda\theta_{00}e^{-\lambda z^0_{0,L}}c\frac{dz^1_{0,L}}{d\tilde{q}} = 0$$

$$\text{(C.2.13)} \quad \frac{d^2\mathcal{L}}{dn^0_{0,L}d\tilde{q}} = -\frac{d\xi_{0,L}}{d\tilde{q}}(1 - e^{-\lambda\bar{z}_{0,L}}) - \xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}} + \frac{d\varphi_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^L_{0,L}} = 0,$$

$$\text{where we use } \frac{d\mathcal{L}}{dz^\ell_{0,L}},\ \ell \leq L$$

$$\text{(C.2.14)} \quad \frac{d^2\mathcal{L}}{d\xi_{0,L}d\tilde{q}} = 1 - \frac{dn^0_{0,L}}{d\tilde{q}}(1 - e^{-\lambda\bar{z}_{0,L}}) - n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}} = 0$$

$$\text{(C.2.15)} \quad \frac{d^2\mathcal{L}}{d\varphi_{0,L}d\tilde{q}} = \frac{dn^0_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^L_{0,L}} - n^0_{0,L}\theta_{00}\lambda e^{-\lambda z^L_{0,L}}\frac{dz^L_{0,L}}{d\tilde{q}} = 0$$

Proposition 1, part a). Define as auxiliary function:

$$f_{0,L} = \begin{cases} 1 & \text{for } L = 0 \\ 1 - \frac{\frac{dz^{L-1}_{0,L}}{d\tilde{q}}}{\frac{dz^L_{0,L}}{d\tilde{q}}} & \text{for } L \geq 1 \end{cases}$$

In the following, we use the property $f_{0,L} > 0$.

- For $L = 1$: $f_{0,1} = 1 - \theta_{00}e^{-\lambda z^0_{0,1}} > 0$. We obtain the expression by using $\partial\mathcal{L}/\partial z^0_{0,L}$.
- For $L = 2$: $f_{0,2} = 1 - {dz^1_{0,2}/d\tilde{q}}/{dz^2_{0,2}/d\tilde{q}} > 0$ if $dz^2_{0,2}/d\tilde{q} > dz^1_{0,2}/d\tilde{q}$. This is the case if $e^{\lambda(z^1_{0,2}-z^0_{0,2})}(1-\theta_{00}e^{-\lambda z^0_{0,2}}) > 1$. $e^{\lambda(z^1_{0,2}-z^0_{0,2})} \geq \lambda z^1_{0,2} + \lambda/c$ by $z^2_{0,2} \geq z^1_{0,2}$ and $1-\theta_{00}e^{-\lambda z^0_{0,2}} \geq (\lambda z^0_{0,2}+\lambda/c-1)/(\lambda z^0_{0,2}+\lambda/c)$ by $z^1_{0,2} \geq z^0_{0,2}$, so the left side of the inequality exceeds $\lambda z^1_{0,2}+\lambda/c / \lambda z^0_{0,2}+\lambda/c \cdot (\lambda z^0_{0,2} + \lambda/c - 1) > 1$ (cf. Assumption 1).

To show: *The knowledge of the CEO $\bar{z}_{0,L}$ increases with output $\tilde{q}$.*
We solve equation (C.2.9) for ${d\bar{z}_{0,L}}/{d\tilde{q}}$ after substituting for ${dn^0_{0,L}}/{d\tilde{q}}$ and ${d\xi_{0,L}}/{d\tilde{q}}$.

1. From equation (C.2.14):

$$\frac{dn^0_{0,L}}{d\tilde{q}} = \frac{1 - n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}}{1 - e^{-\lambda\bar{z}_{0,L}}} \tag{C.2.16}$$

2. From equation (C.2.13):

$$\frac{d\xi_{0,L}}{d\tilde{q}} = \frac{\frac{d\varphi_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^L_{0,L}} - \xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}}{1 - e^{-\lambda\bar{z}_{0,L}}}$$

From equation (C.2.10):

$$\frac{d\varphi_{0,L}}{d\tilde{q}} = \varphi_{0,L}\lambda f_{0,L}\frac{dz^L_{0,L}}{d\tilde{q}} \tag{C.2.17}$$

where the derivation is straightforward for $L = 0$. For $L \geq 1$, we employ the definition of $\varphi_{0,L}$ to simplify the equation.
From equation (C.2.15):

$$\frac{dz^L_{0,L}}{d\tilde{q}} = \frac{1}{\lambda n^0_{0,L}}\frac{dn^0_{0,L}}{d\tilde{q}} \tag{C.2.18}$$

Using equations (C.2.16)-(C.2.18) to substitute for ${d\varphi_{0,L}}/{d\tilde{q}}$ in ${d\xi_{0,L}}/{d\tilde{q}}$, we obtain

$$\begin{aligned}\frac{d\xi_{0,L}}{d\tilde{q}} = \frac{1}{1 - e^{-\lambda\bar{z}_{0,L}}}\Bigg(&\varphi_{0,L}f_{0,L}\theta_{00}e^{-\lambda z^L_{0,L}}\frac{1 - n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}}{n^0_{0,L}(1 - e^{-\lambda\bar{z}_{0,L}})}\\ &- \xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}\Bigg)\end{aligned} \tag{C.2.19}$$

3. Substituting equations (C.2.16) and (C.2.19) into equation (C.2.9) and solving for ${d\bar{z}_{0,L}}/{d\tilde{q}}$ yields:

$$\begin{aligned}\frac{d\bar{z}_{0,L}}{d\tilde{q}} &= \frac{1}{n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}}\\ &\times \frac{\xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}} + \frac{\lambda e^{-\lambda\bar{z}_{0,L}}}{1-e^{-\lambda\bar{z}_{0,L}}}\theta_{00}e^{-\lambda z^L_{0,L}}\varphi_{0,L}f_{0,L}}{\xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}} + \frac{\lambda e^{-\lambda\bar{z}_{0,L}}}{1-e^{-\lambda\bar{z}_{0,L}}}\theta_{00}e^{-\lambda z^L_{0,L}}\varphi_{0,L}f_{0,L} + \lambda\xi_{0,L}} > 0.\end{aligned} \tag{C.2.20}$$

□

To show: *The number of employees $n^\ell_{0,L}$ at all below-CEO layers $\ell \leq L$ increases with output $\tilde{q}$.*

1. $\ell = 0$: Equation (C.2.16) implies that ${dn^0_{0,L}}/{d\tilde{q}} > 0$ if ${d\bar{z}_{0,L}}/{d\tilde{q}} < {1}/{\lambda n^0_{0,L}e^{-\lambda\bar{z}_{0,L}}}$. The condition holds, because the second factor in equation (C.2.20) is below one.

2. $\ell = L$, $L \geq 1$: Using constraint (5) and equation (C.2.18):

$$\frac{dn^L_{0,L}}{d\tilde{q}} = \frac{dn^0_{0,L}}{d\tilde{q}} \theta_{00} e^{-\lambda z^{L-1}_{0,L}} f_{0,L} > 0$$

3. $\ell = 1$, $L = 2$: Using constraint (5) and equations (C.2.11), (C.2.12), (C.2.16):

$$\frac{dn^1_{0,2}}{d\tilde{q}} = \frac{dn^0_{0,2}}{d\tilde{q}} \theta_{00} e^{-\lambda z^0_{0,2}} \left(1 - \frac{\theta_{00} e^{-\lambda z^1_{0,2}}}{1 - \theta_{00} e^{-\lambda z^0_{0,2}}}\right) > 0$$

by $\theta_{00} e^{-\lambda z^1_{0,2}} \leq \theta_{00} e^{-\lambda z^0_{0,2}} = \dfrac{c}{\lambda(1 + c z^1_{0,2})} < \dfrac{1}{2}$ by Assumption 1 ($\lambda > 2c$) □

To show: *The knowledge $z^\ell_{0,L}$ of the employees at all below-CEO layers $\ell \leq L$ increase with output $\tilde{q}$.*

1. $\ell = L$: $dz^L_{0,L}/d\tilde{q} > 0$ by equation (C.2.18) and $dn^0_{0,L}/d\tilde{q} > 0$.
2. $\ell = 1$, $L = 2$: From equations (C.2.11) and (C.2.12):

$$\frac{dz^1_{0,2}}{d\tilde{q}} = \frac{dz^2_{0,2}}{d\tilde{q}} \frac{e^{-\lambda z^1_{0,2}}}{e^{-\lambda z^0_{0,2}}(1 - \theta_{00} e^{-\lambda z^0_{0,2}})} > 0.$$

   We use $\partial\mathcal{L}/\partial z^1_{0,2}$ and $\partial\mathcal{L}/\partial z^0_{0,2}$ to simplify the expression.
3. $\ell = 0$, $L > 0$: From equation (C.2.12):

□
$$\frac{dz^0_{0,L}}{d\tilde{q}} = \frac{dz^1_{0,L}}{d\tilde{q}} \frac{c}{\lambda(1 + c z^1_{0,L})} > 0.$$

To show: *The managerial span of control $n^{\ell-1}_{0,L}/n^\ell_{0,L}$ at all managerial layers $1 \leq \ell \leq L+1$ increases with output $\tilde{q}$.*

1. $\ell = 1$, $L \geq 1$: The span of control is $\theta_{00}^{-1} e^{\lambda z^0_{0,L}}$, which increases by $dz^0_{0,L}/d\tilde{q} > 0$.
2. $\ell = 2$, $L = 2$: The span of control is $e^{\lambda(z^1_{0,2} - z^0_{0,2})}$, which increases by $dz^1_{0,2}/d\tilde{q} > dz^0_{0,2}/d\tilde{q}$ from equation (C.2.12) and $\partial\mathcal{L}/\partial z^0_{0,2}$.
3. $\ell = L + 1$: The CEO's span of control is $n^L_{0,L}$, which increases by $dn^L_{0,L}/d\tilde{q} > 0$. □

To show: *The marginal benefit of CEO time $\varphi_{0,L}$ increases with output $\tilde{q}$.*
Follows from equation (C.2.17) by $f_{0,L} > 0$ and $dz^L_{0,L}/d\tilde{q} > 0$. □

Proposition 1, part b).

To show: *The cost function $C_{0,L}(\tilde{q})$ strictly increases with output $\tilde{q}$.*
Follows from $\partial C_{0,L}(\tilde{q})/\partial\tilde{q} = \xi_{0,L} > 0$. □

To show: *The marginal costs increase with output for $\tilde{q} \geq \hat{q}^L$, with $\hat{q}^0 = 0$.*
From equations (C.2.19) and (C.2.20):

$$\frac{d\xi_{0,L}}{d\tilde{q}} > 0 \quad \text{if} \quad \varphi_{0,L} f_{0,L} \theta_{00} e^{-\lambda z^L_{0,L}} > \xi_{0,L} e^{-\lambda \bar{z}_{0,L}}$$

For $L = 0$, this condition holds $\forall \tilde{q}$ by ${}^{w_0 c}/_{\lambda} > {}^{w_0 c}/_{\lambda n^0_{0,0}}$.

For $L > 0$, the condition is equivalent to $e^{\lambda(z^L_{0,L} - z^{L-1}_{0,L})} > f^{-1}_{0,L}$. It holds for $\tilde{q} \geq \hat{q}^L$. Importantly, it holds at the MES, where $e^{\lambda(z^L_{0,L} - z^{L-1}_{0,L})} = \lambda \bar{z}_{0,L} + {}^{\lambda}/_{c}$. $\square$

$$
\begin{aligned}
\text{For } L = 1: \qquad & \lambda \bar{z}_{0,1} + \frac{\lambda}{c} > f^{-1}_{0,1} = {}^{\lambda z^1_{0,1} + \frac{\lambda}{c}} / {}_{\lambda z^1_{0,1} + \frac{\lambda}{c} - 1} \\
\text{by} \qquad & \lambda \bar{z}_{0,1} + \frac{\lambda}{c} > \lambda z^1_{0,1} + \frac{\lambda}{c} \quad \text{and} \quad \lambda z^1_{0,1} + \frac{\lambda}{c} - 1 > 1 \quad \text{by Assumption 1} \\
\text{For } L = 2: \qquad & \lambda \bar{z}_{0,2} + \frac{\lambda}{c} > f^{-1}_{0,2} = {}^{e^{-\lambda z^0_{0,2}}(1 - \theta_{00} e^{-\lambda z^0_{0,2}})} / {}_{e^{-\lambda z^0_{0,2}}(1 - \theta_{00} e^{-\lambda z^0_{0,2}}) - e^{-\lambda z^1_{0,2}}} \\
\text{by} \qquad & \lambda \bar{z}_{0,2} + \frac{\lambda}{c} > 2 > f^{-1}_{0,2} \\
\text{because } & e^{\lambda(z^1_{0,2} - z^0_{0,2})}(1 - \theta_{00} e^{-\lambda z^0_{0,2}}) = \frac{\lambda(1 + c z^2_{0,2})}{\lambda(1 + c z^1_{0,2})} \frac{\lambda(1 + c z^1_{0,2}) - c}{c} > 2 \\
& \text{as } \lambda > 2c \text{ by Assumption 1 and } \lambda z^1_{0,2} > 1
\end{aligned}
$$

To show: *The average cost function $AC_{0,L}(\tilde{q})$ reaches a minimum at $\tilde{q}^{*L}$ where it intersects with the marginal cost function, and converges to infinity for $\tilde{q} \to 0$ and $\tilde{q} \to \infty$.*

$$
\begin{aligned}
AC_{0,L}(\tilde{q}) &= \frac{C_{0,L}(\tilde{q})}{\tilde{q}} \\
\Rightarrow \quad \frac{dAC_{0,L}(\tilde{q})}{d\tilde{q}} &= \frac{1}{\tilde{q}} \left(\xi_{0,L} - AC_{0,L} \right) = 0 \text{ if } \xi_{0,L} = AC_{0,L} \\
\frac{d^2 AC_{0,L}(\tilde{q})}{d\tilde{q}^2} &= -\frac{2}{\tilde{q}^2} \left(\xi_{0,L} - AC_{0,L} \right) + \frac{1}{\tilde{q}} \frac{d\xi_{0,L}}{d\tilde{q}} = \frac{1}{\tilde{q}} \frac{d\xi_{0,L}}{d\tilde{q}} > 0 \text{ if } \xi_{0,L} = AC_{0,L} \\
\lim_{\tilde{q} \to 0} AC_{0,L}(\tilde{q}) &= \infty \text{ because } C_{0,L}(\tilde{q}) \geq w_0 \text{ and } C_{0,L}(\tilde{q}) < \infty \text{ for } \tilde{q} \to 0. \\
\lim_{\tilde{q} \to \infty} AC_{0,L}(\tilde{q}) &= \infty \text{ because } \lim_{\tilde{q} \to \infty} \xi_{0,L} = \infty \text{ by l'Hôpital's rule.}
\end{aligned}
$$
$\square$

C.2.4. The optimal number of layers

We follow Caliendo and Rossi-Hansberg (2012, p. 1454 et seqq.). We show that the average cost function has a unique minimum at the minimum efficient scale $\tilde{q}^{*L}$ for a given number of below-CEO layers L and that the minimum average costs decrease with L.

That the minimum efficient scale $\tilde{q}^{*L}$ increases with the number of below-CEO layers L follows from Caliendo and Rossi-Hansberg (2012, p. 1456-8).

To prove uniqueness of the minimum efficient scale (MES), we show that there exists a unique cut-point of the first order conditions (FOCs) and the respective condition for the MES. We focus on positive solutions for the knowledge levels.

The FOCs (7), (8) and (C.2.1) define the optimal knowledge levels recursively:

$$
\lambda \bar{z}_{0,L} - \lambda z^L_{0,L} = \ln \left(\lambda z^0_{0,L} + \frac{\lambda}{c} + 1 + \theta_{00} \sum_{\ell=0}^{L} e^{-\lambda z^\ell_{0,L}} \right) - \ln \theta_{00}
$$

$$\lambda z^1_{0,L} - \lambda z^0_{0,L} = \ln\left(\lambda z^2_{0,L} + \frac{\lambda}{c}\right) \qquad \text{for } L > 1$$

$$\lambda z^0_{0,L} = \ln\left(\lambda z^1_{0,L} + \frac{\lambda}{c}\right) + \ln\theta_{00} \qquad \text{for } L > 0$$

At the MES, $AC_{0,L} = \xi_{0,L}$:

$$\lambda z^0_{0,0} = \ln\left(\lambda \bar{z}_{0,0} + \frac{\lambda}{c}\right) + \ln\theta_{00} \qquad \text{for } L = 0$$

$$\lambda z^L_{0,L} - \lambda z^{L-1}_{0,L} = \ln\left(\lambda \bar{z}_{0,L} + \frac{\lambda}{c}\right) \qquad \text{for } L > 0$$

Both the FOCs and the conditions for the MES define $z^L_{0,L}$ as (implicit) functions of $\bar{z}_{0,L}$. The FOCs have a positive root:

$$z^L_{0,L} = 0: \quad \lambda\bar{z}_{0,L} \geq \ln\left(\frac{\lambda}{c} + 1 + \theta_{00}\right) - \ln\theta_{00} > 0$$

The conditions for the MES have a positive intercept:

$$L = 0,\ \bar{z}_{0,0} = 0: \quad \lambda z^0_{0,0} = \ln\left(\frac{\lambda\theta_{00}}{c}\right) > 0 \quad \text{by Assumption 1}$$

$$L > 0,\ \bar{z}_{0,L} = 0: \quad \lambda z^L_{0,L} - \lambda z^{L-1}_{0,L} = \ln\left(\frac{\lambda}{c}\right) > \ln\left(\frac{\lambda\theta_{00}}{c}\right)$$

Both the conditions for the MES and the f.o.c.s are strictly increasing:

$$\text{MES}: \quad \frac{dz^L_{0,L}}{d\bar{z}_{0,L}} = \frac{1}{\lambda\bar{z}_{0,L} + \frac{\lambda}{c}}\frac{1}{f_{0,L}} \qquad > 0$$

$$\text{FOC},\ L = 0: \quad \frac{dz^0_{0,0}}{d\bar{z}_{0,0}} = \frac{\lambda z^0_{0,0} + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z^0_{0,0}}}{\lambda z^0_{0,0} + \frac{\lambda}{c} + 1 + f_{0,0}} \qquad > 0$$

$$\text{FOC},\ L = 1: \quad \frac{dz^1_{0,1}}{d\bar{z}_{0,1}} = \frac{\lambda z^0_{0,1} + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z^0_{0,1}} + \theta_{00}e^{-\lambda z^1_{0,1}}}{\lambda z^0_{0,1} + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z^0_{0,1}} + \theta_{00}e^{-\lambda z^0_{0,1}} f_{0,1}} \qquad > 0$$

$$\text{FOC},\ L = 2: \quad \frac{dz^2_{0,2}}{d\bar{z}_{0,2}} = \frac{\lambda z^0_{0,2} + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z^\ell_{0,2}} + \theta_{00}e^{-\lambda z^2_{0,2}}}{\lambda z^0_{0,2} + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z^\ell_{0,2}} + \theta_{00}e^{-\lambda z^1_{0,2}} f_{0,2}} \qquad > 0$$

where $f_{0,L}$ is defined in section C.2.3.

The slope of the conditions for the MES decreases continuously with $\bar{z}_{0,L}$ from a value smaller than 1 with $\lim_{\bar{z}_{0,L}\to\infty} dz^L_{0,L}/d\bar{z}_{0,L} = 0$. The slope of the FOCs is close to 1 with $\lim_{\bar{z}_{0,L}\to\infty} dz^L_{0,L}/d\bar{z}_{0,L} = 1$. Thus, for a given number of layers L, there exists a unique cut-point of the FOC and the condition for the MES.

Proposition 5 (see below) implies that the minimum average costs (MAC) of a single-establishment organization with L below-CEO layers cannot exceed those of an organization with $L-1$ below-CEO layers, i.e. $MAC_{0,L-1} \geq MAC_{0,L}$. $\square$

C.3. Multi-establishment firm organization

C.3.1. Results for $L_j \geq 2$ layers of middle managers

Constraints (17) and (18) apply to layer $\ell = 2, ..., L_j$. The endogenous variables are stated in section C.3.2.

For the comparative statics, we restrict our attention to firms with $L_j + 1 \leq 3$ managerial layers in line with sections II and III.

Proposition C.3. *The results in Proposition 3 apply to a multi-establishment firm with $L_j = 2$ layers of middle managers at the establishment, the headquarters, or both.*

Proof. See Appendix C.3.4. □

Proposition C.4. *The results in Proposition 4 apply to a multi-establishment firm with $L_0 = 2$ layers of middle managers at the headquarters and $L_1 < 2$ layers of middle managers at the establishment. They also apply to a multi-establishment firm with $L_0 \leq 2$ layers of middle managers at the headquarters and $L_1 = 2$ layers of middle managers at the establishment if the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high. Concerning the CEO span of control $\sum_{j=0}^{1} n_{j,\omega}^{L_j}$, the results apply if $L_0 = 0$ and $L_1 = 2$, if $L_0 = 2$ and $L_1 = 0$ and $w_0 > w_1$, and if $L_j = 2$, $L_k > 0$, $k \neq j$ and the share of CEO time $s_{j,\omega}$ and the ratio w_0/w_1 are sufficiently high.*

Proof. See Appendix C.3.5. □

C.3.2. Lagrangian equation and first order conditions

Firm-level: CEO knowledge, allocation of CEO time and output

$$\begin{aligned}
\mathcal{L} = & \sum_{j=0}^{1} C_{j,\omega}(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}) + \left(1 - \sum_{j=0}^{1} s_{j,\omega}\right) w_0(1 + c\bar{z}_{0,\omega}) \\
& + \bar{\kappa}_{0,\omega}\left(\sum_{j=0}^{1} s_{j,\omega} - 1\right) - \sum_{j=0}^{1} \kappa_{j,\omega} s_{j,\omega} - \eta_{0,\omega}\bar{z}_{0,\omega} + \bar{\phi}_{0,\omega}\left(\sum_{j=0}^{1} \tilde{q}_j - \sum_{j=0}^{1} q_{j,\omega}\right) - \sum_{j=0}^{1} \phi_{j,\omega} q_{j,\omega} \\
& + \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\underline{\phi}_{0,\omega}(\tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega}) - q_{0,\omega}) + \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{1,\omega}(\tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega}) - q_{1,\omega})
\end{aligned}$$

First order conditions:

$$\text{(C.3.1)} \qquad \frac{\partial \mathcal{L}}{\partial q_{0,\omega}} = \frac{\partial C_{0,\omega}}{\partial q_{0,\omega}} - \bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\underline{\phi}_{0,\omega} - \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{1,\omega}\tau - \phi_{0,\omega} = 0$$

$$\text{(C.3.2)} \qquad \frac{\partial \mathcal{L}}{\partial q_{1,\omega}} = \frac{\partial C_{1,\omega}}{\partial q_{1,\omega}} - \bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\underline{\phi}_{0,\omega}\tau - \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{1,\omega} - \phi_{1,\omega} = 0$$

$$\text{(C.3.3)} \qquad \frac{\partial \mathcal{L}}{\partial s_{j,\omega}} = \frac{\partial C_{j,\omega}}{\partial s_{j,\omega}} - w_0(1 + c\bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} - \kappa_{j,\omega} = 0$$

$$\text{(C.3.4)} \qquad \frac{\partial \mathcal{L}}{\partial \bar{z}_{0,\omega}} = \textstyle\sum_{j=0}^{1} \frac{\partial C_{j,\omega}}{\partial \bar{z}_{0,\omega}} + w_0 c(1 - s_{0,\omega} - s_{1,\omega}) - \eta_{0,\omega} = 0$$

$$\text{(C.3.5)} \qquad \frac{\partial \mathcal{L}}{\partial \bar{\kappa}_{0,\omega}} = s_{0,\omega} + s_{1,\omega} - 1 = 0$$

$$\text{(C.3.6)} \qquad \frac{\partial \mathcal{L}}{\partial \bar{\phi}_{0,\omega}} = \sum_{j=0}^{1} \tilde{q}_j - \sum_{j=0}^{1} q_{j,\omega} = 0$$

$$\text{(C.3.7)} \qquad \frac{\partial \mathcal{L}}{\partial \underline{\phi}_{0,\omega}} = \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)(\tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega}) - q_{0,\omega}) = 0$$

$$\text{(C.3.8)} \qquad \frac{\partial \mathcal{L}}{\partial \underline{\phi}_{1,\omega}} = \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)(\tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega}) - q_{1,\omega}) = 0$$

Establishment-level: Number and knowledge of employees. We use equation (17), which is binding in optimum, to substitute for $n^{\ell}_{j,\omega}$, $L_j \geq \ell > 0$.

$$\begin{aligned}
\mathcal{L} = {} & n^0_{j,\omega} w_j \left(1 + c z^0_{j,\omega}\right) + \mathbb{1}(L_j \geq 1) n^0_{j,\omega} \sum_{\ell=1}^{L_j} \theta_{jj} e^{-\lambda z^{\ell-1}_{j,\omega}} w_j \left(1 + c z^{\ell}_{j,\omega}\right) + s_{j,\omega} w_0 \left(1 + c\bar{z}_{0,\omega}\right) \\
& + \xi_{j,\omega} \left(q_{j,\omega} - n^0_{j,\omega}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right)\right) + \varphi_{j,\omega}\left(n^0_{j,\omega} \theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - s_{j,\omega}\right) \\
& + \bar{\eta}^{L_j+1}_{j,\omega}(z^{L_j}_{j,\omega} - \bar{z}_{0,\omega}) + \mathbb{1}(L_j \geq 1) \sum_{\ell=1}^{L_j} \bar{\eta}^{\ell}_{j,\omega}(z^{\ell-1}_{j,\omega} - z^{\ell}_{j,\omega}) - \bar{\eta}^0_{j,\omega} z^0_{j,\omega} - \eta^0_{j,\omega} n^0_{j,\omega}
\end{aligned}$$

$$\text{(C.3.9)} \qquad \frac{\partial \mathcal{L}}{\partial z^{L_j}_{j,\omega}} \begin{cases} \overset{L_j=0}{=} n^0_{j,\omega}\left(w_j c - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^0_{j,\omega}}\right) + \bar{\eta}^1_{j,\omega} - \bar{\eta}^0_{j,\omega} = 0 \\ \overset{L_j \geq 1}{=} n^0_{j,\omega}\left(w_j c \theta_{jj} e^{-\lambda z^{L_j-1}_{j,\omega}} - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}}\right) + \bar{\eta}^{L_j+1}_{j,\omega} - \bar{\eta}^{L_j}_{j,\omega} = 0 \end{cases}$$

$$\text{(C.3.10)} \qquad \frac{\partial \mathcal{L}}{\partial z^{\ell}_{j,\omega}} = n^0_{j,\omega} w_j \left(c\theta_{jj} e^{-\lambda z^{\ell-1}_{j,\omega}} - \lambda \theta_{jj} e^{-\lambda z^{\ell}_{j,\omega}}(1 + c z^{\ell+1}_{j,\omega})\right) + \bar{\eta}^{\ell+1}_{j,\omega} - \bar{\eta}^{\ell}_{j,\omega} = 0$$

for $L_j > \ell > 0$, $L_j \geq 2$

$$\text{(C.3.11)} \qquad \frac{\partial \mathcal{L}}{\partial z^0_{j,\omega}} \overset{L_j \geq 1}{=} n^0_{j,\omega} w_j \left(c - \lambda\theta_{jj} e^{-\lambda z^0_{j,\omega}}(1 + c z^1_{j,\omega})\right) + \bar{\eta}^1_{j,\omega} - \bar{\eta}^0_{j,\omega} = 0$$

$$\begin{aligned}
& \frac{\partial \mathcal{L}}{\partial n^0_{j,\omega}} = w_j \left(1 + c z^0_{j,\omega} + \mathbb{1}(L_j \geq 1) \sum_{\ell=1}^{L_j} \theta_{jj} e^{-\lambda z^{\ell-1}_{j,\omega}} \left(1 + c z^{\ell}_{j,\omega}\right)\right) \\
\text{(C.3.12)} \qquad & - \xi_{j,\omega}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right) + \varphi_{j,\omega}\theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - \eta^0_{j,\omega} = 0
\end{aligned}$$

$$\text{(C.3.13)} \qquad \frac{\partial \mathcal{L}}{\partial \xi_{j,\omega}} = q_{j,\omega} - n^0_{j,\omega}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right) = 0$$

$$\text{(C.3.14)} \qquad \frac{\partial \mathcal{L}}{\partial \varphi_{j,\omega}} = n^0_{j,\omega}\theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - s_{j,\omega} = 0$$

Endogenous variables:

$$\begin{aligned}
e^{\lambda z^{L_j}_{j,\omega}} &= \frac{q_{j,\omega}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{\theta_{j0}}{s_{j,\omega}} \\
e^{\lambda\left(z^{\ell-1}_{j,\omega} - z^{\ell-2}_{j,\omega}\right)} &= \left(1 + c z^{\ell}_{j,\omega}\right)\frac{\lambda}{c} \quad \forall \ell = 2, ..., L_j,\ L_j \geq 2 \\
e^{\lambda z^0_{j,\omega}} &= \left(1 + c z^1_{j,\omega}\right)\frac{\lambda}{c}\theta_{jj} \quad \text{for } L_j \geq 1 \\
\xi_{j,\omega} &= \frac{w_j\left(1 + c z^0_{j,\omega} + \frac{c}{\lambda} + \mathbb{1}(L_j \geq 1)\theta_{jj}\frac{c}{\lambda}\sum_{\ell=1}^{L_j} e^{-\lambda z^{\ell-1}_{j,\omega}}\right)}{1 - e^{-\lambda \bar{z}_{0,\omega}}}
\end{aligned}$$

$$\varphi_{j,\omega} = \frac{w_j c}{\lambda \theta_{j0}} \theta_{jj} e^{\lambda\left(z_{j,\omega}^{L_j} - z_{j,\omega}^{L_j - 1}\right)} \quad \text{for } L_j \geq 1, \quad \varphi_{j,\omega} = \frac{w_j c}{\lambda \theta_{j0}} e^{\lambda z_{j,\omega}^0} \quad \text{for } L_j = 0$$

C.3.3. Proposition 2: Allocation of output and CEO time

To show: *In optimum,* $\varphi_{0,\omega} = \varphi_{1,\omega}$.

From equation (C.3.3): If $\kappa_{j,\omega} = 0 \, \forall j$, i.e., the CEO spends positive time $s_{0,\omega}, s_{1,\omega} > 0$ at both locations,

$$\begin{aligned} \bar{\kappa}_{0,\omega} &= w_0(1 + c\bar{z}_{0,\omega}) - \frac{\partial C}{\partial s_{0,\omega}} \\ &= w_0(1 + c\bar{z}_{0,\omega}) - \frac{\partial C}{\partial s_{1,\omega}} \\ \Rightarrow \quad & \varphi_{0,\omega} = \varphi_{1,\omega} \end{aligned}$$

□

To show: *In optimum,*

(1) $\xi_{0,\omega} \leq \tau\xi_{1,\omega} \wedge \xi_{1,\omega} \leq \tau\xi_{0,\omega}$ *if* $q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0$.

(2) $\tau\xi_{0,\omega} = \xi_{1,\omega}$ *if* $q_{1,\omega} < \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega})$

(3) $\xi_{0,\omega} = \tau\xi_{1,\omega}$ *if* $q_{0,\omega} < \tilde{q}_0 \wedge q_{1,\omega} = \tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega})$

(1) Assume $q_{0,\omega} = \tilde{q}_0 \wedge q_{1,\omega} = \tilde{q}_1$. $\bar{\phi}_{0,\omega}$, $\underline{\phi}_{j,\omega} \geq 0$, $\phi_{j,\omega} = 0 \, \forall j$. From equations (C.3.1) and (C.3.2), with $\partial C / \partial q_{j,\omega} = \xi_{j,\omega}$:

$$\begin{aligned} \underline{\phi}_{0,\omega} &= \frac{1}{\tau^2 - 1}\left(\tau\xi_{1,\omega} - \xi_{0,\omega} - (\tau - 1)\bar{\phi}_{0,\omega}\right) \\ \text{By } \bar{\phi}_{0,\omega}, \underline{\phi}_{0,\omega} &\geq 0 : \\ \xi_{0,\omega} &\leq \tau\xi_{1,\omega} - (\tau - 1)\bar{\phi}_{0,\omega} \leq \tau\xi_{1,\omega} \\ \underline{\phi}_{1,\omega} &= \frac{1}{\tau^2 - 1}\left(\tau\xi_{0,\omega} - \xi_{1,\omega} - (\tau - 1)\bar{\phi}_{0,\omega}\right) \\ \text{By } \bar{\phi}_{0,\omega}, \underline{\phi}_{1,\omega} &\geq 0 : \\ \xi_{1,\omega} &\leq \tau\xi_{0,\omega} - (\tau - 1)\bar{\phi}_{0,\omega} \leq \tau\xi_{0,\omega} \end{aligned}$$

(2, 3) Assume $q_{j,\omega} > 0 \, \forall j$, $q_{0,\omega} > \tilde{q}_0 \wedge q_{1,\omega} < \tilde{q}_1$. $\phi_{j,\omega} = 0 \, \forall j$, and $\underline{\phi}_{0,\omega} \geq 0$.
$\bar{\phi}_{0,\omega} = 0$ by

$$\begin{aligned} q_{0,\omega} + q_{1,\omega} &= \tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega}) + q_{1,\omega} \\ &= \tilde{q}_0 + \tilde{q}_1 + (\tau - 1)(\tilde{q}_1 - q_{1,\omega}) \\ &> \tilde{q}_0 + \tilde{q}_1, \text{ i.e., the constraint is slack.} \end{aligned}$$

From equation (C.3.1), $\underline{\phi}_{0,\omega} = \xi_{0,\omega}$. Inserting this into equation (C.3.2) yields

$$\xi_{0,\omega} = \tau\xi_{1,\omega}.$$

The analogous result holds if $q_{j,\omega} > 0 \, \forall j$, $q_{1,\omega} > \tilde{q}_1 \wedge q_{0,\omega} < \tilde{q}_0$.
Note: If $\exists j$ s.t. $\phi_{j,\omega} > 0$, $\xi_{j,\omega} > \tau\xi_{k,\omega} \; k \neq j$ at $q_{j,\omega} = 0$. To see this, consider $\phi_{1,\omega} > 0$. In this case, $\underline{\phi}_{0,\omega} = \xi_{0,\omega}$ and $\xi_{1,\omega} = \tau\underline{\phi}_{0,\omega} + \phi_{1,\omega}$, so $\xi_{1,\omega} = \tau\xi_{0,\omega} + \phi_{1,\omega} > \tau\xi_{0,\omega}$. □

To show: *If $\tau = 1$, in optimum, $\xi_{0,\omega} = \xi_{1,\omega}$.*
From equations (C.3.1) and (C.3.2):

$$\begin{aligned} 0 &= \frac{\partial C}{\partial q_{0,\omega}} - \bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\underline{\phi}_{0,\omega} - \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{1,\omega} \\ &= \frac{\partial C}{\partial q_{1,\omega}} - \bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\underline{\phi}_{0,\omega} - \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{0,\omega} \end{aligned}$$

□ $$\Rightarrow \quad \xi_{0,\omega} = \xi_{1,\omega}$$

C.3.4. Proposition 3, C.3: Comparative statics with respect to $\tilde{q}_j$

We differentiate the first order conditions with respect to output and solve the resulting system of linear equations. We assume that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$ and $q_{j,\omega}, s_{j,\omega} > 0\,\forall j$. The second order conditions are, with $k, j \in \{0, 1\}$:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_j} = \frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \frac{d\bar{\phi}_{0,\omega}}{d\tilde{q}_j} - \frac{d\underline{\phi}_{0,\omega}}{d\tilde{q}_j} - \tau\frac{d\underline{\phi}_{1,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.15}$$

$$\frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_j} = \frac{d\xi_{1,\omega}}{d\tilde{q}_j} - \frac{d\bar{\phi}_{0,\omega}}{d\tilde{q}_j} - \tau\frac{d\underline{\phi}_{0,\omega}}{d\tilde{q}_j} - \frac{d\underline{\phi}_{1,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.16}$$

$$\frac{d^2\mathcal{L}}{ds_{k,\omega}d\tilde{q}_j} = -\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} + \frac{d\bar{\kappa}_{0,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.17}$$

$$\begin{aligned} \frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}_j} &= -\sum_{k=0}^{1}\frac{d\xi_{k,\omega}}{d\tilde{q}_j}n^0_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}} - \sum_{k=0}^{1}\xi_{k,\omega}\frac{dn^0_{k,\omega}}{d\tilde{q}_j}\lambda e^{-\lambda\bar{z}_{0,\omega}} \\ &\quad + \sum_{k=0}^{1}\xi_{k,\omega}n^0_{k,\omega}\lambda^2 e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = 0 \end{aligned} \tag{C.3.18}$$

$$\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\tilde{q}_j} = \frac{ds_{0,\omega}}{d\tilde{q}_j} + \frac{ds_{1,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.19}$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_j} = 1 - \frac{dq_{0,\omega}}{d\tilde{q}_j} - \frac{dq_{1,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.20}$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_j} = \mathbb{1}(q_{1,\omega} \leq \tilde{q}_1)\left(\mathbb{1}(j=0)1 + \mathbb{1}(j=1)\tau - \frac{dq_{0,\omega}}{d\tilde{q}_j} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_j}\right) = 0 \tag{C.3.21}$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{1,\omega}d\tilde{q}_j} = \mathbb{1}(q_{0,\omega} \leq \tilde{q}_0)\left(\mathbb{1}(j=1)1 + \mathbb{1}(j=0)\tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_j} - \frac{dq_{1,\omega}}{d\tilde{q}_j}\right) = 0 \tag{C.3.22}$$

$$\frac{d^2\mathcal{L}}{dz^{L_k}_{k,\omega}d\tilde{q}_j} \begin{cases} \overset{L_k=0}{=} -\frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^0_{k,\omega}} + \varphi_{k,\omega}\theta_{k0}\lambda e^{-\lambda z^0_{k,\omega}}\frac{dz^0_{k,\omega}}{d\tilde{q}_j} = 0 \\ \overset{L_k\geq 1}{=} -w_k c\theta_{kk}e^{-\lambda z^{L_k-1}_{k,\omega}}\frac{dz^{L_k-1}_{k,\omega}}{d\tilde{q}_j} - \frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} \\ \qquad +\varphi_{k,\omega}\theta_{k0}\lambda e^{-\lambda z^{L_k}_{k,\omega}}\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} = 0 \end{cases} \tag{C.3.23}$$

$$\begin{aligned} \frac{d^2\mathcal{L}}{dz^{\ell}_{k,\omega}d\tilde{q}_j} &= -ce^{-\lambda z^{\ell-1}_{k,\omega}}\frac{dz^{\ell-1}_{k,\omega}}{d\tilde{q}_j} + \lambda e^{-\lambda z^{\ell}_{k,\omega}}\frac{dz^{\ell}_{k,\omega}}{d\tilde{q}_j}(1 + cz^{\ell+1}_{k,\omega}) - e^{-\lambda z^{\ell}_{k,\omega}}c\frac{dz^{\ell+1}_{k,\omega}}{d\tilde{q}_j} = 0 \\ &\text{for } L_k > \ell > 0,\ L_k \geq 2 \end{aligned} \tag{C.3.24}$$

$$\text{(C.3.25)}\qquad \frac{d^2\mathcal{L}}{dz^0_{k,\omega}d\tilde{q}_j} \overset{L_k\geq 1}{=} \lambda^2\theta_{kk}e^{-\lambda z^0_{k,\omega}}\frac{dz^0_{k,\omega}}{d\tilde{q}_j}(1+cz^1_{k,\omega}) - \lambda\theta_{kk}e^{-\lambda z^0_{k,\omega}}c\frac{dz^1_{k,\omega}}{d\tilde{q}_j} = 0$$

$$\text{(C.3.26)}\qquad \frac{d^2\mathcal{L}}{dn^0_{k,\omega}d\tilde{q}_j} = -\frac{d\xi_{k,\omega}}{d\tilde{q}_j}(1-e^{-\lambda\bar{z}_{0,\omega}}) - \xi_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} + \frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} = 0$$

where we substitute $\frac{d\mathcal{L}}{dz^\ell_{k,\omega}}$, $\ell \leq L_k$

$$\text{(C.3.27)}\qquad \frac{d^2\mathcal{L}}{d\xi_{k,\omega}d\tilde{q}_j} = \frac{dq_{k,\omega}}{d\tilde{q}_j} - \frac{dn^0_{k,\omega}}{d\tilde{q}_j}(1-e^{-\lambda\bar{z}_{0,\omega}}) - n^0_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = 0$$

$$\text{(C.3.28)}\qquad \frac{d^2\mathcal{L}}{d\varphi_{k,\omega}d\tilde{q}_j} = \frac{dn^0_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} - n^0_{k,\omega}\theta_{k0}\lambda e^{-\lambda z^{L_k}_{k,\omega}}\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} - \frac{ds_{k,\omega}}{d\tilde{q}_j} = 0$$

Auxiliary results.

1. *Local production $q_{j,\omega}$ varies one-to-one with local output $\tilde{q}_j$, but does not vary with output at the other location $\tilde{q}_k$, $k \neq j$.*
   From equation (C.3.22):

$$\frac{dq_{1,\omega}}{d\tilde{q}_j} = \mathbb{1}(j=1)1 + \mathbb{1}(j=0)\tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_j}$$

   Into equation (C.3.21):

$$\mathbb{1}(j=0)(1-\tau^2) = (1-\tau^2)\frac{dq_{0,\omega}}{d\tilde{q}_j} \quad\Rightarrow\quad \frac{q_{0,\omega}}{d\tilde{q}_j} = 1 \text{ for } j=0,\ \frac{q_{0,\omega}}{d\tilde{q}_j} = 0 \text{ for } j=1$$

   Together with equation (C.3.20), this yields:

$$\text{(C.3.29)}\qquad \frac{q_{j,\omega}}{d\tilde{q}_j} = 1; \quad \frac{q_{k,\omega}}{d\tilde{q}_j} = 0,\ k\neq j$$

2. *An increase of the share of CEO time $s_{j,\omega}$ with output at location j leads to an equal decrease of the share of CEO time $s_{k,\omega}$ at the other location $k \neq j$.*
   From equation (C.3.19):

$$\text{(C.3.30)}\qquad \frac{ds_{1,\omega}}{d\tilde{q}_j} = -\frac{ds_{0,\omega}}{d\tilde{q}_j}$$

3. *Changes in the marginal benefit of CEO time $\varphi_{k,\omega}$ with local output $\tilde{q}_j$ are equal across locations.*
   From equation (C.3.17) for $k = 0, 1$:

$$\text{(C.3.31)}\qquad \frac{d\varphi_{0,\omega}}{d\tilde{q}_j} = \frac{d\varphi_{1,\omega}}{d\tilde{q}_j}$$

Proposition 3, part a). Define as auxiliary function

$$f_{k,\omega} = \begin{cases} 1 & \text{for } L_k = 0 \\ 1 - \frac{dz^{L_k-1}_{k,\omega}}{d\tilde{q}} / \frac{dz^{L_k}_{k,\omega}}{d\tilde{q}} & \text{for } L_k \geq 1 \end{cases}$$

$f_{k,\omega} > 0$ (see section C.2.3).

To show: *CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_j$.*
We solve equation (C.3.18) for $d\bar{z}_{0,\omega}/d\tilde{q}_j$ after substituting for $dn^0_{k,\omega}/d\tilde{q}_j$ and $d\xi_{k,\omega}/d\tilde{q}_j$.

1. From equation (C.3.27):

$$\frac{dn^0_{k,\omega}}{d\tilde{q}_j} = \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\left(\frac{q_{k,\omega}}{d\tilde{q}_j} - n^0_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j}\right) \tag{C.3.32}$$

2. From equation (C.3.26):

$$\frac{d\xi_{k,\omega}}{d\tilde{q}_j} = \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\left(\frac{d\varphi_{0,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} - \xi_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j}\right)$$

From equation (C.3.23):

$$\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = \varphi_{k,\omega}\lambda f_{k,\omega}\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} \tag{C.3.33}$$

where the derivation is straightforward for $L_k = 0$. For $L_k \geq 1$, we employ the definition of $\varphi_{k,\omega}$ to simplify the equation.
From equation (C.3.28):

$$\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} = -\frac{1}{\lambda s_{k,\omega}}\frac{ds_{k,\omega}}{d\tilde{q}_j} + \frac{1}{\lambda n^0_{k,\omega}}\frac{dn^0_{k,\omega}}{d\tilde{q}_j} \tag{C.3.34}$$

We use equations (C.3.30), (C.3.31), (C.3.33) and (C.3.34) to find $\frac{ds_{k,\omega}}{d\tilde{q}_j}$, $k = 0, 1$:

$$\frac{ds_{0,\omega}}{d\tilde{q}_j} = \frac{\frac{f_{0,\omega}}{n^0_{0,\omega}}\frac{dn^0_{0,\omega}}{d\tilde{q}_j} - \frac{f_{1,\omega}}{n^0_{1,\omega}}\frac{dn^0_{1,\omega}}{d\tilde{q}_j}}{\sum_{k=0}^{1}\frac{f_{k,\omega}}{s_{k,\omega}}} \tag{C.3.35}$$

$$\frac{ds_{1,\omega}}{d\tilde{q}_j} = \frac{\frac{f_{1,\omega}}{n^0_{1,\omega}}\frac{dn^0_{1,\omega}}{d\tilde{q}_j} - \frac{f_{0,\omega}}{n^0_{0,\omega}}\frac{dn^0_{0,\omega}}{d\tilde{q}_j}}{\sum_{k=0}^{1}\frac{f_{k,\omega}}{s_{k,\omega}}} \tag{C.3.36}$$

Substituting for $ds_{k,\omega}/d\tilde{q}_j$ and subsequently $dn^0_{k,\omega}/d\tilde{q}_j$ in equation (C.3.34) with $k = 0$, we obtain:

$$\begin{aligned}\frac{dz^{L_k}_{0,\omega}}{d\tilde{q}_j} &= \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{f_{1,\omega}}{\sum_{k=0}^{1}\frac{f_{k,\omega}}{s_{k,\omega}}}\frac{1}{\lambda s_{0,\omega}s_{1,\omega}}\left(-\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} + \sum_{k=0}^{1}\frac{s_{k,\omega}}{n^0_{k,\omega}}\frac{dq_{k,\omega}}{d\tilde{q}_j}\right)\\ &= \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{f_{1,\omega}}{\sum_{k=0}^{1}\frac{f_{k,\omega}}{s_{k,\omega}}}\frac{1}{\lambda s_{0,\omega}s_{1,\omega}}\left(-\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} + \frac{s_{j,\omega}}{n^0_{j,\omega}}\right)\end{aligned} \tag{C.3.37}$$

where the second equality follows from equation (C.3.29).

From equation (C.3.38):

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}_j} = \varphi_{0,\omega} \frac{\theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} - \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j}}{(1 - e^{-\lambda \bar{z}_{0,\omega}}) \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}}} \tag{C.3.38}$$

We use equation (C.3.38) to substitute for ${d\varphi_{0,\omega}}/{d\tilde{q}_j}$ in ${d\xi_{k,\omega}}/{d\tilde{q}_j}$:

$$\begin{aligned}\frac{d\xi_{k,\omega}}{d\tilde{q}_j} = \frac{1}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \Bigg(& \frac{1}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{\varphi_{0,\omega} \theta_{k0} e^{-\lambda z_{k,\omega}^{L_k}}}{\sum_{m=0}^{1} \frac{s_{m,\omega}}{f_{m,\omega}}} \left(\theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} - \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} \right) \\ & - \xi_{k,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} \Bigg)\end{aligned} \tag{C.3.39}$$

3. Substituting equations (C.3.32) (accounting for C.3.29) and (C.3.39) into equation (C.3.18) and solving for ${d\bar{z}_{0,\omega}}/{d\tilde{q}_j}$ yields:

$$\begin{aligned}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = & \frac{1}{n_{j,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}}} \\ & \times \frac{\xi_{j,\omega} n_{j,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}} \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}} + s_{j,\omega} \varphi_{0,\omega} \frac{\lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}}}{\lambda (1 + e^{-\lambda \bar{z}_{0,\omega}}) \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}} \sum_{k=0}^{1} \xi_{k,\omega} n_{k,\omega}^0 + \varphi_{0,\omega} \frac{\lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}}} > 0\end{aligned} \tag{C.3.40}$$

□

To show: *Higher local output $\tilde{q}_j$ increases (decreases) the number of production workers $n_{j,\omega}^0$ ($n_{k,\omega}^0$) at location j (location $k \neq j$).*

Follows from equation (C.3.32) by ${d\bar{z}_{0,\omega}}/{d\tilde{q}_j} > 0$ and ${dq_{k,\omega}}/{d\tilde{q}_j} = 0$ for location k, and ${d\bar{z}_{0,\omega}}/{d\tilde{q}_j} < {1}/{n_{j,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}}}$ and ${dq_{j,\omega}}/{d\tilde{q}_j} = 1$ for location j. □

To show: *Higher local output $\tilde{q}_j$ increases (decreases) the share of CEO time $s_{j,\omega}$ ($s_{k,\omega}$) at location j (location $k \neq j$).*

Follows from equations (C.3.35) and (C.3.36) by ${dn_{j,\omega}^0}/{d\tilde{q}_j} > 0$ and ${dn_{k,\omega}^0}/{d\tilde{q}_j} < 0$. □

Proposition 3, part b).

To show: *The knowledge of the employees at all below-CEO layers $z_{k,\omega}^{\ell}$, $\ell \leq L_k$, $k = 0, 1$, the below-CEO managerial span of control ${n_{j,\omega}^{\ell-1}}/{n_{j,\omega}^{\ell}}$ and the marginal benefit of CEO time $\varphi_{k,\omega}$ increase with local output $\tilde{q}_j$ if the CEO spends a sufficient share of time on location j.*

From equation (C.3.38): ${d\varphi_{0,\omega}}/{d\tilde{q}_j} > 0$ if $\theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} > \lambda e^{-\lambda \bar{z}_{0,\omega}} {d\bar{z}_{0,\omega}}/{d\tilde{q}_j}$

This is the case if

$$s_{j,\omega} \geq \underbrace{\frac{1}{1 + e^{\lambda \bar{z}_{0,\omega}}}}_{\leq 0.5} \underbrace{\frac{\xi_{j,\omega} n_{j,\omega}^0}{\sum_{k=0}^{1} \xi_{k,\omega} n_{k,\omega}^0}}_{<1}$$

The positive impact of higher local output on knowledge follows from equation (C.3.33):

$$\frac{dz_{k,\omega}^{L_k}}{d\tilde{q}_j} \overset{L_k=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\tilde{q}_j} \qquad \frac{dz_{k,\omega}^{L_k}}{d\tilde{q}_j} f_{k,\omega} \overset{L_k\geq 1}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\tilde{q}_j}$$

Knowledge at lower layers depends on knowledge at higher layers as in the single-establishment firm (see section C.2.3).

The positive impact on the below-CEO managerial span of control follows from:

$$n_{j,\omega}^{\ell-1}/n_{j,\omega}^{\ell} = \begin{cases} e^{\lambda z_{j,\omega}^0}/\theta_{jj} & \text{for } \ell = 1 \\ e^{\lambda(z_{j,\omega}^1 - z_{j,\omega}^0)} & \text{for } \ell = 2 \end{cases}$$

and ${dz_{j,\omega}^0}/{d\tilde{q}_j} > 0$ and ${dz_{j,\omega}^1}/{d\tilde{q}_j} - {dz_{j,\omega}^0}/{d\tilde{q}_j} > 0$. □

To show: *The marginal production costs $\xi_{k,\omega}$ increase with output $\tilde{q}_j$ if CEO knowledge is sufficiently high.*

From equation (C.3.39):

$$\frac{d\xi_{k,\omega}}{d\tilde{q}_j} \geq 0 \quad \text{if} \quad \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} \leq \frac{1}{n_{j,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}}} \times \frac{\varphi_{0,\omega}\theta_{k0} e^{-\lambda z_{k,\omega}^{L_k}} s_{j,\omega}}{\varphi_{0,\omega}\theta_{k0} e^{-\lambda z_{k,\omega}^{L_k}} + \xi_{k,\omega}(1 - e^{-\lambda \bar{z}_{0,\omega}}) \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}}}$$

For $j = k$, this is the case if:

$$e^{\lambda \bar{z}_{0,\omega}} + 1 \geq \frac{\xi_{j,\omega}\tilde{q}_j}{\sum_{k=0}^{1} \xi_{k,\omega}\tilde{q}_k} \frac{1}{s_{j,\omega}} \left(2 + \frac{\xi_{j,\omega}\tilde{q}_j}{\varphi_{0,\omega} s_{j,\omega}} \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}} \right)$$ □

For $j \neq k$, this is the case if:

$$e^{\lambda \bar{z}_{0,\omega}} + 1 \geq \frac{\xi_{j,\omega}\tilde{q}_j}{\sum_{k=0}^{1} \xi_{k,\omega}\tilde{q}_k} \frac{1}{s_{j,\omega}} \left(1 + \frac{\xi_{k,\omega}\tilde{q}_k s_{j,\omega}}{\xi_{j,\omega}\tilde{q}_j s_{k,\omega}} + \frac{\xi_{k,\omega}\tilde{q}_k}{\varphi_{0,\omega} s_{k,\omega}} \sum_{k=0}^{1} \frac{s_{k,\omega}}{f_{k,\omega}} \right)$$ □

C.3.5. Proposition 4, C.4: Comparative statics with respect to θ_{10}

We differentiate the first order conditions with respect to the helping costs θ_{10} and solve the resulting system of linear equations. We assume that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$ and $q_{j,\omega}$, $s_{j,\omega} > 0\,\forall j$. The second order conditions are, with $j \in \{0, 1\}$:

$$\text{(C.3.41)} \qquad \frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\bar{\phi}_{0,\omega}}{d\theta_{10}} - \frac{d\underline{\phi}_{0,\omega}}{d\theta_{10}} - \tau\frac{d\underline{\phi}_{1,\omega}}{d\theta_{10}} = 0$$

$$\text{(C.3.42)} \qquad \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \frac{d\xi_{1,\omega}}{d\theta_{10}} - \frac{d\bar{\phi}_{0,\omega}}{d\theta_{10}} - \tau\frac{d\underline{\phi}_{0,\omega}}{d\theta_{10}} - \frac{d\underline{\phi}_{1,\omega}}{d\theta_{10}} = 0$$

$$\text{(C.3.43)} \qquad \frac{d^2\mathcal{L}}{ds_{j,\omega}d\theta_{10}} = -\frac{d\varphi_{j,\omega}}{d\theta_{10}} + \frac{d\bar{\kappa}_{0,\omega}}{d\theta_{10}} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\theta_{10}} = -\sum_{j=0}^{1} \frac{d\xi_{j,\omega}}{d\theta_{10}} n_{j,\omega}^0 \lambda e^{-\lambda\bar{z}_{0,\omega}} - \sum_{j=0}^{1} \xi_{j,\omega} \frac{dn_{j,\omega}^0}{d\theta_{10}} \lambda e^{-\lambda\bar{z}_{0,\omega}}$$

$$+\sum_{j=0}^{1}\xi_{j,\omega}n^0_{j,\omega}\lambda^2 e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}=0 \tag{C.3.44}$$

$$\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\theta_{10}}=\frac{ds_{0,\omega}}{d\theta_{10}}+\frac{ds_{1,\omega}}{d\theta_{10}}=0 \tag{C.3.45}$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}}=-\frac{dq_{0,\omega}}{d\theta_{10}}-\frac{dq_{1,\omega}}{d\theta_{10}}=0 \tag{C.3.46}$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\theta_{10}}=\mathbb{1}(q_{1,\omega}\le\tilde{q}_1)\left(-\frac{dq_{0,\omega}}{d\theta_{10}}-\tau\frac{dq_{1,\omega}}{d\theta_{10}}\right)=0 \tag{C.3.47}$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{1,\omega}d\theta_{10}}=\mathbb{1}(q_{0,\omega}\le\tilde{q}_0)\left(-\tau\frac{dq_{0,\omega}}{d\theta_{10}}+\frac{dq_{1,\omega}}{d\theta_{10}}\right)=0 \tag{C.3.48}$$

$$\frac{d^2\mathcal{L}}{dz^{L_0}_{0,\omega}d\theta_{10}}\begin{cases}\overset{L_0=0}{=}-\frac{d\varphi_{0,\omega}}{d\theta_{10}}+\varphi_{0,\omega}\lambda\frac{dz^0_{0,\omega}}{d\theta_{10}}=0\\ \overset{L_0\ge1}{=}-w_0ce^{-\lambda z^{L_0-1}_{0,\omega}}\frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}}-\frac{d\varphi_{0,\omega}}{d\theta_{10}}e^{-\lambda z^{L_0}_{0,\omega}}+\varphi_{0,\omega}\lambda e^{-\lambda z^{L_0}_{0,\omega}}\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}}=0\end{cases} \tag{C.3.49}$$

where we cancel $\lambda\theta_{00}$ and, in the first equation, $e^{-\lambda z^0_{0,\omega}}$.

$$\frac{d^2\mathcal{L}}{dz^{L_1}_{1,\omega}d\theta_{10}}\begin{cases}\overset{L_1=0}{=}-\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}+\varphi_{1,\omega}\theta_{10}\lambda\frac{dz^0_{1,\omega}}{d\theta_{10}}-\varphi_{1,\omega}=0\\ \overset{L_1\ge1}{=}-w_1c\theta_{11}e^{-\lambda z^{L_1-1}_{1,\omega}}\frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}}-\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}\\ +\varphi_{1,\omega}\theta_{10}\lambda e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}}-\varphi_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}}=0\end{cases} \tag{C.3.50}$$

where we cancel λ and, in the first equation, $e^{-\lambda z^0_{1,\omega}}$.

$$\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}d\theta_{10}}=-ce^{-\lambda z^{\ell-1}_{j,\omega}}\frac{dz^{\ell-1}_{j,\omega}}{d\theta_{10}}+\lambda e^{-\lambda z^{\ell}_{j,\omega}}\frac{dz^{\ell}_{j,\omega}}{d\theta_{10}}(1+cz^{\ell+1}_{j,\omega})-e^{-\lambda z^{\ell}_{j,\omega}}c\frac{dz^{\ell+1}_{j,\omega}}{d\theta_{10}}=0 \tag{C.3.51}$$

for $L_j>\ell>0$, $L_j\ge 2$

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\theta_{10}}\overset{L_j\ge1}{=}\lambda^2\theta_{jj}e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\theta_{10}}(1+cz^1_{j,\omega})-\lambda\theta_{jj}e^{-\lambda z^0_{j,\omega}}c\frac{dz^1_{j,\omega}}{d\theta_{10}}=0 \tag{C.3.52}$$

$$\frac{d^2\mathcal{L}}{dn^0_{0,\omega}d\theta_{10}}=-\frac{d\xi_{0,\omega}}{d\theta_{10}}(1-e^{-\lambda\bar{z}_{0,\omega}})-\xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}+\frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}}=0, \tag{C.3.53}$$

where we substitute $d\mathcal{L}/dz^{\ell}_{0,\omega}$, $\ell\le L_0$.

$$\frac{d^2\mathcal{L}}{dn^0_{1,\omega}d\theta_{10}}=-\frac{d\xi_{1,\omega}}{d\theta_{10}}(1-e^{-\lambda\bar{z}_{0,\omega}})-\xi_{1,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}+\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}$$

$$+\varphi_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}}=0,\ \text{where we substitute } d\mathcal{L}/dz^{\ell}_{1,\omega},\ \ell\le L_1. \tag{C.3.54}$$

$$\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}=\frac{dq_{j,\omega}}{d\theta_{10}}-\frac{dn^0_{j,\omega}}{d\theta_{10}}(1-e^{-\lambda\bar{z}_{0,\omega}})-n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}=0 \tag{C.3.55}$$

$$\frac{d^2\mathcal{L}}{d\varphi_{0,\omega}d\theta_{10}}=\frac{dn^0_{0,\omega}}{d\theta_{10}}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}}-n^0_{0,\omega}\theta_{00}\lambda e^{-\lambda z^{L_0}_{0,\omega}}\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}}-\frac{ds_{0,\omega}}{d\theta_{10}}=0 \tag{C.3.56}$$

$$\frac{d^2\mathcal{L}}{d\varphi_{1,\omega}d\theta_{10}}=\frac{dn^0_{1,\omega}}{d\theta_{10}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}-n^0_{1,\omega}\theta_{10}\lambda e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}}-\frac{ds_{1,\omega}}{d\theta_{10}}+n^0_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}}=0 \tag{C.3.57}$$

Auxiliary results.

1. *Local production $q_{j,\omega}$ does not vary with the helping costs θ_{10}.*
   From equation (C.3.48):
$$\frac{dq_{1,\omega}}{d\theta_{10}} = -\tau \frac{dq_{0,\omega}}{d\theta_{10}}$$
   From equation (C.3.47):
$$\frac{dq_{1,\omega}}{d\theta_{10}} = -\frac{1}{\tau} \frac{dq_{0,\omega}}{d\theta_{10}}$$
   From equation (C.3.46):
$$\frac{dq_{1,\omega}}{d\theta_{10}} = -\frac{dq_{0,\omega}}{d\theta_{10}}$$
   This implies:
$$\frac{dq_{0,\omega}}{d\theta_{10}} = \frac{1}{\tau} \frac{dq_{0,\omega}}{d\theta_{10}} = \tau \frac{dq_{0,\omega}}{d\theta_{10}}$$
$$\Rightarrow \quad \frac{dq_{0,\omega}}{d\theta_{10}} = 0 \text{ by } \tau > 1; \quad \frac{dq_{1,\omega}}{d\theta_{10}} = 0 \tag{C.3.58}$$

2. *An increase of the share of CEO time $s_{j,\omega}$ with the helping costs θ_{10} at location j leads to an equal decrease of the share of CEO time $s_{k,\omega}$ at the other location $k \neq j$.*
   From equation (C.3.45):
$$\frac{ds_{1,\omega}}{d\theta_{10}} = -\frac{ds_{0,\omega}}{d\theta_{10}} \tag{C.3.59}$$

3. *Changes in the marginal benefit of CEO time $\varphi_{k,\omega}$ with the helping costs θ_{10} are equal across locations.*
   From equation (C.3.43) for $k = 0, 1$:
$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{d\varphi_{1,\omega}}{d\theta_{10}} \tag{C.3.60}$$

Proposition 4, part a). Note: while some results hold if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high, the comparative statics results concerning below-CEO knowledge $z^{\ell}_{1,\omega}$, the below-CEO span of control and the marginal production costs $\xi_{1,\omega}$ hold in general.

To show: *CEO knowledge $\bar{z}_{0,\omega}$ increases with the helping costs θ_{10} if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high.*

We solve equation (C.3.44) for ${d\bar{z}_{0,\omega}}/{d\theta_{10}}$ after substituting for ${dn^0_{j,\omega}}/{d\theta_{10}}$ and ${d\xi_{j,\omega}}/{d\theta_{10}}$.

1. From equation (C.3.55) with equation (C.3.58):
$$\frac{dn^0_{j,\omega}}{d\theta_{10}} = -\frac{1}{1 - e^{-\lambda \bar{z}_{0,\omega}}} n^0_{j,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} \tag{C.3.61}$$

2. From equations (C.3.53, C.3.54):

$$\frac{d\xi_{j,\omega}}{d\theta_{10}} = \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\left(\frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{j0}e^{-\lambda z_{j,\omega}^{L_j}} - \xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \mathbb{1}(j=1)\varphi_{1,\omega}e^{-\lambda z_{1,\omega}^{L_1}}\right)$$

From equation (C.3.49):

$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \varphi_{0,\omega}\lambda f_{0,\omega}\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} \tag{C.3.62}$$

and from equation (C.3.50):

$$\frac{d\varphi_{1,\omega}}{d\theta_{10}} = \varphi_{0,\omega}\lambda f_{1,\omega}\frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} - \frac{1}{\theta_{10}}\varphi_{0,\omega} \tag{C.3.63}$$

where the derivation is straightforward for $L_j = 0$. For $L_j \geq 1$, we employ the definition of $\varphi_{0,\omega}$ to simplify the equations.
From equations (C.3.56, C.3.57):

$$\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} = -\frac{1}{\lambda s_{0,\omega}}\frac{ds_{0,\omega}}{d\theta_{10}} + \frac{1}{\lambda n_{0,\omega}^0}\frac{dn_{0,\omega}^0}{d\theta_{10}} \tag{C.3.64}$$

$$\frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} = -\frac{1}{\lambda s_{1,\omega}}\frac{ds_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda n_{1,\omega}^0}\frac{dn_{1,\omega}^0}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}} \tag{C.3.65}$$

We use equations (C.3.59), (C.3.60), (C.3.62), (C.3.63), (C.3.64) and (C.3.65) to find $\frac{ds_{j,\omega}}{d\theta_{10}}$:

$$\frac{ds_{0,\omega}}{d\theta_{10}} = -s_{0,\omega}\left(\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \frac{1}{f_{0,\omega}\varphi_{0,\omega}}\frac{d\varphi_{0,\omega}}{d\theta_{10}}\right) \tag{C.3.66}$$

$$\frac{ds_{1,\omega}}{d\theta_{10}} = -s_{1,\omega}\left(\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \frac{1}{f_{1,\omega}\varphi_{0,\omega}}\frac{d\varphi_{0,\omega}}{d\theta_{10}} + \frac{1-f_{1,\omega}}{\theta_{10}f_{1,\omega}}\right) \tag{C.3.67}$$

Using equation (C.3.59), we obtain:

$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \varphi_{0,\omega}\frac{\frac{s_{1,\omega}}{\theta_{10}}\frac{f_{1,\omega}-1}{f_{1,\omega}} - \frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}}{\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}}} \tag{C.3.68}$$

We use equation (C.3.68) to substitute for ${d\varphi_{0,\omega}}/{d\theta_{10}}$ in ${d\xi_{j,\omega}}/{d\theta_{10}}$:

$$\begin{aligned}\frac{d\xi_{j,\omega}}{d\theta_{10}} = \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}&\left(\frac{\varphi_{0,\omega}\theta_{j0}e^{-\lambda z_{j,\omega}^{L_j}}}{\sum_{k=0}^{1}\frac{s_{k,\omega}}{f_{k,\omega}}}\left(\frac{s_{1,\omega}}{\theta_{10}}\frac{f_{1,\omega}-1}{f_{1,\omega}} - \frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}\right)\right.\\ &\left. - \xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \mathbb{1}(j=1)\varphi_{1,\omega}e^{-\lambda z_{1,\omega}^{L_1}}\right)\end{aligned} \tag{C.3.69}$$

3. Substituting equations (C.3.61) and (C.3.69) into equation (C.3.44) and solving for ${d\bar{z}_{0,\omega}}/{d\theta_{10}}$

yields:

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \varphi_{0,\omega} n^0_{1,\omega} e^{-\lambda z^{L_1}_{1,\omega}}$$

$$\times \frac{\sum_{j=0}^{1} \frac{s_{j,\omega}}{f_{j,\omega}} - \frac{1-f_{1,\omega}}{f_{1,\omega}}}{\lambda(1+e^{-\lambda \bar{z}_{0,\omega}}) \sum_{j=0}^{1} \frac{s_{j,\omega}}{f_{j,\omega}} \sum_{j=0}^{1} \xi_{j,\omega} n^0_{j,\omega} + \varphi_{0,\omega} \frac{\lambda e^{-\lambda \bar{z}_{0,\omega}}}{1-e^{-\lambda \bar{z}_{0,\omega}}}} \tag{C.3.70}$$

The denominator is positive, so the sign depends on the numerator.

- For $L_1 = 0$, $f_{1,\omega}(\varphi_{0,\omega}) = 1$, so $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.
- For $L_1 = 1$, $\sum_{j=0}^{1} \frac{s_{j,\omega}}{f_{j,\omega}} > 1 > \frac{1-f_{1,\omega}}{f_{1,\omega}}$. $1/f_{1,\omega} - 1 < 1$ by $1/f_{1,\omega} = 1/1-\theta_{11} e^{-\lambda z^0_{1,\omega}} < \lambda/c+\lambda z^0_{1,\omega} / \lambda/c+\lambda z^0_{1,\omega}-1 < 2$ by Assumption 1, so $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.
- For $L_1 = 2$, $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ if $s_{1,\omega} > 1 - f_{1,\omega}$ (sufficient, not necessary). □

To show: *The knowledge of the employees at all below-CEO layers $z^\ell_{1,\omega}$, $\ell \leq L_1$, at the establishment increases with the helping costs θ_{10}.*

From equation (C.3.63):

$$\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda} \frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$$f_{1,\omega} \frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} \overset{L_1\geq 1}{=} \frac{1}{\varphi_{1,\omega}\lambda} \frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$\frac{dz^\ell_{1,\omega}}{d\theta_{10}} > 0$ because $\frac{\varphi_{1,\omega}}{\theta_{10}} > -\frac{d\varphi_{1,\omega}}{d\theta_{10}}$ by $s_{0,\omega}/\theta_{10} f_{0,\omega} + s_{1,\omega}/\theta_{10} > \lambda e^{-\lambda \bar{z}_{0,\omega}} / (1-e^{-\lambda \bar{z}_{0,\omega}}) d\bar{z}_{0,\omega}/d\theta_{10}$. Knowledge at lower layers depends on knowledge at higher layers as in the single-establishment firm (see section C.2.3). □

To show: *The managerial span of control $n^{\ell-1}_{1,\omega}/n^\ell_{1,\omega}$, $1 \leq \ell \leq L_1$, at the establishment increases with the helping costs θ_{10}.*

Follows from

$$n^{\ell-1}_{1,\omega}/n^\ell_{1,\omega} = \begin{cases} e^{\lambda z^0_{1,\omega}}/\theta_{11} & \text{for } \ell = 1 \\ e^{\lambda(z^1_{1,\omega} - z^0_{1,\omega})} & \text{for } \ell = 2 \end{cases}$$

and $dz^0_{1,\omega}/d\theta_{10} > 0$ and $dz^1_{1,\omega}/\theta_{10} - dz^0_{1,\omega}/\theta_{10} > 0$. □

To show: *The marginal production costs $\xi_{1,\omega}$ at the establishment increase with the helping costs θ_{10}.*

Follows from equation (C.3.69) because $\varphi_{1,\omega} e^{-\lambda z^{L_1}_{1,\omega}} > \xi_{1,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{0,\omega}}{d\theta_{10}} \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}$, which results after substituting for $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}$ and $\frac{d\varphi_{0,\omega}}{d\theta_{10}}$ by $s_{1,\omega} > 0 \geq \frac{f_{1,\omega}-1}{f_{1,\omega}}$. □

To show: *The total number of production workers $\sum_{j=0}^{1} n^0_{j,\omega}$ as well as the number of production workers $n^0_{1,\omega}$ at the establishment decrease with the helping costs θ_{10} if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high.*

Follows from equation (C.3.61) and $d\bar{z}_{0,\omega}/d\theta_{10} > 0$. □.

To show: *The share of CEO time $s_{1,\omega}$ at the establishment decreases with the helping costs θ_{10} if $\exists j$ s.t. $L_j > 0$; it is constant otherwise. This result holds if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high.*

From equations (C.3.67) and (C.3.68):

$$\frac{ds_{1,\omega}}{d\theta_{10}} = \frac{s_{0,\omega}s_{1,\omega}}{\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}}}\frac{1-f_{1,\omega}}{\theta_{10}f_{0,\omega}f_{1,\omega}}\left(\frac{\lambda\theta_{10}e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{f_{0,\omega}-f_{1,\omega}}{1-f_{1,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}-1\right)$$

- For $L_0 = L_1 = 0$: $f_{j,\omega} = 1\forall j$, so ${ds_{1,\omega}}/{d\theta_{10}} = 0$.
- For $\exists j$ s.t. $L_j > 0$:

$$\begin{aligned}&\frac{\lambda\theta_{10}e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{f_{0,\omega}-f_{1,\omega}}{1-f_{1,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}\\ =&\frac{\varphi_{0,\omega}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}s_{1,\omega}\frac{f_{0,\omega}-f_{1,\omega}}{1-f_{1,\omega}}\left(\frac{s_{0,\omega}}{f_{0,\omega}}-\frac{s_{0,\omega}}{f_{1,\omega}}+1\right)}{\varphi_{0,\omega}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}+\lambda(1+e^{-\lambda\bar{z}_{0,\omega}})\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}}\sum_{j=0}^{1}\xi_{j,\omega}n_{j,\omega}^{0}}\end{aligned}$$

  For $f_{0,\omega} \geq f_{1,\omega}$, ${ds_{1,\omega}}/{d\theta_{10}} < 0$ because $f_{0,\omega} - f_{1,\omega} \leq 1 - f_{1,\omega}$ and ${s_{0,\omega}}/{f_{0,\omega}} \leq {s_{0,\omega}}/{f_{1,\omega}}$, so the expression is smaller than 1.

  For $f_{0,\omega} < f_{1,\omega}$, ${ds_{1,\omega}}/{d\theta_{10}} < 0$ because the expression is smaller than 0. □

Proposition 4, part b). Note: while most results hold if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high, the comparative statics results concerning the marginal production costs $\xi_{0,\omega}$ hold in general.

To show: *The knowledge of the employees at all below-CEO layers $z_{0,\omega}^{\ell}$, $\ell \leq L_0$, at the headquarters decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_{1,\omega}$.*

From equation (C.3.62):

$$\begin{aligned}\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} &\overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}}\\ f_{0,\omega}\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} &\overset{L_0\geq 1}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}}\end{aligned}$$

$\frac{dz_{0,\omega}^{\ell}}{d\theta_{10}} < 0$ follows from $\frac{d\varphi_{0,\omega}}{d\theta_{10}} < 0$ by equation (C.3.68) due to $f_{1,\omega} \leq 1$ and ${d\bar{z}_{0,\omega}}/{d\theta_{10}} > 0$. Knowledge at lower layers depends on knowledge at higher layers as in the single-establishment firm (see section C.2.3). □

To show: *The managerial span of control ${n_{0,\omega}^{\ell-1}}/{n_{0,\omega}^{\ell}}$, $1 \leq \ell \leq L_0$, at the headquarters decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_{1,\omega}$.*

Follows from

$${n_{0,\omega}^{\ell-1}}/{n_{0,\omega}^{\ell}} = \begin{cases} e^{\lambda z_{0,\omega}^{0}}/\theta_{00} & \text{for } \ell = 1\\ e^{\lambda(z_{0,\omega}^{1}-z_{0,\omega}^{0})} & \text{for } \ell = 2\end{cases}$$

and ${dz^{L_0}_{0,\omega}}/{d\theta_{10}} < 0$ and ${dz^1_{0,\omega}}/{\theta_{10}} - {dz^0_{0,\omega}}/{\theta_{10}} > 0$. $\square$

To show: *The marginal production costs $\xi_{0,\omega}$ at the establishment decrease with the helping costs θ_{10}.*
Follows from equation (C.3.69) after substituting for ${d\bar{z}_{0,\omega}}/{d\theta_{10}}$ due to $f_{1,\omega} \leq 1$. $\square$

To show: *The number of production workers $n^0_{0,\omega}$ at the headquarters decreases with the helping costs θ_{10} if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high.*
Follows from equation (C.3.61) and ${d\bar{z}_{0,\omega}}/{d\theta_{10}} > 0$. $\square$

To show: *The headquarter share of CEO time $s_{0,\omega}$ increases if $\exists j$ s.t. $L_j > 0$; it is constant otherwise. This result holds if $L_1 \leq 1$ or if $L_1 = 2$ the share of CEO time $s_{1,\omega}$ at the establishment is sufficiently high.*
Follows from ${ds_{1,\omega}}/{d\theta_{10}} < 0$ and equation (C.3.59). $\square$

To show: *The CEO span of control $\sum_{j=0}^{1} n^{L_j}_{j,\omega}$ decreases with the helping costs θ_{10} for $L_j \leq 1\,\forall j$ if $\exists j$ s.t. $L_j = 0$ or the ratio ${w_0}/{w_1}$ is sufficiently high. The CEO span of control also decreases with the helping costs θ_{10} if $L_0 = 0$ and $L_1 = 2$, if $L_0 = 2$ and $L_1 = 0$ and $w_0 > w_1$, and if $L_j = 2$, $L_k > 0$, $k \neq j$ and the share of CEO time $s_{j,\omega}$ and the ratio ${w_0}/{w_1}$ are sufficiently high.*

$$\sum_{j=0}^{1} n^{L_j}_{j,\omega} = \begin{cases} n^0_{0,\omega} + n^0_{1,\omega} & \text{if } L_0 = L_1 = 0 \\ n^0_{0,\omega} + n^0_{1,\omega}\theta_{11}e^{-\lambda z^{L_1-1}_{1,\omega}} & \text{if } L_0 = 0 < L_1 \\ n^0_{0,\omega}\theta_{00}e^{-\lambda z^{L_0-1}_{0,\omega}} + n^0_{1,\omega} & \text{if } L_0 > L_1 = 0 \\ n^0_{0,\omega}\theta_{00}e^{-\lambda z^{L_0-1}_{0,\omega}} + n^0_{1,\omega}\theta_{11}e^{-\lambda z^{L_1-1}_{1,\omega}} & \text{if } L_0 > 0, L_1 > 0 \end{cases}$$

$\Rightarrow$ For $L_0 = L_1 = 0$: ${d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}/{d\theta_{10}} < 0$ by ${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0, 1$.

For $L_0 = 0 < L_1$: ${d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}/{d\theta_{10}} < 0$ by ${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0, 1$ and ${dz^{L_1-1}_{1,\omega}}/{d\theta_{10}} > 0$.

For $L_0 > L_1 = 0$:

$$\frac{d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}{d\theta_{10}} = \frac{dn^0_{0,\omega}}{d\theta_{10}}\theta_{00}e^{\lambda z^{L_0-1}_{0,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^{L_0-1}_{0,\omega}}\frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}} + \frac{dn^0_{1,\omega}}{d\theta_{10}}$$

${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0, 1$, but ${dz^{L_0-1}_{0,\omega}}/{d\theta_{10}} < 0$.
For $L_0 = 1$, $-{dn^0_{0,\omega}}/{d\theta_{10}}\theta_{00}e^{\lambda z^{L_0-1}_{0,\omega}} > -n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^{L_0-1}_{0,\omega}}\,{dz^{L_0-1}_{0,\omega}}/{d\theta_{10}}$ by $\sum_{j=0}^{1} {s_{j,\omega}}/{f_{j,\omega}} > 1$ and Assumption 1.
For $L_0 = 2$, $-{dn^0_{0,\omega}}/{d\theta_{10}}\theta_{00}e^{\lambda z^{L_0-1}_{0,\omega}} - {dn^0_{1,\omega}}/{d\theta_{10}} > -n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^{L_0-1}_{0,\omega}}\,{dz^{L_0-1}_{0,\omega}}/{d\theta_{10}}$ if $w_0 > w_1$ (sufficient, not necessary). This result can be derived by exploiting $\varphi_{0,\omega} = \varphi_{1,\omega}$.

For $L_0 = L_1 = 1$:

$$\frac{d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}{d\theta_{10}} = \sum_{j=0}^{1}\frac{dn^0_{j,\omega}}{d\theta_{10}}\theta_{jj}e^{\lambda z^0_{j,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^0_{0,\omega}}\frac{dz^0_{0,\omega}}{d\theta_{10}} - n^0_{1,\omega}\theta_{11}\lambda e^{\lambda z^0_{1,\omega}}\frac{dz^0_{1,\omega}}{d\theta_{10}}$$

${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0,1$ and ${dz^0_{1,\omega}}/{d\theta_{10}} > 0$, but ${dz^0_{0,\omega}}/{d\theta_{10}} < 0$. The CEO span of control is nevertheless negative if ${w_0}/{w_1} \geq {[1-f_{0,\omega}]}/{[1+f_{0,\omega}s_{1,\omega}/s_{0,\omega}]}$ (sufficient, not necessary).

For $L_0 = 1 < L_1$:

$$\frac{d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}{d\theta_{10}} = \sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{d\theta_{10}} \theta_{jj} e^{\lambda z^{L_j-1}_{j,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^0_{0,\omega}} \frac{dz^0_{0,\omega}}{d\theta_{10}} - n^0_{1,\omega}\theta_{11}\lambda e^{\lambda z^1_{1,\omega}} \frac{dz^1_{1,\omega}}{d\theta_{10}}$$

${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0,1$ and ${dz^1_{1,\omega}}/{d\theta_{10}} > 0$, but ${dz^0_{0,\omega}}/{d\theta_{10}} < 0$. The CEO span of control is nevertheless negative if $s_{1,\omega} > 1 - f_{1,\omega}$ and ${w_0}/{w_1} \geq {[1-f_{0,\omega}]}/{[1+f_{0,\omega}s_{1,\omega}/s_{0,\omega}]}$ (sufficient, not necessary).

For $L_0 > L_1 = 1$:

$$\frac{d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}{d\theta_{10}} = \sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{d\theta_{10}} \theta_{jj} e^{\lambda z^{L_j-1}_{j,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^1_{0,\omega}} \frac{dz^1_{0,\omega}}{d\theta_{10}} - n^0_{1,\omega}\theta_{11}\lambda e^{\lambda z^0_{1,\omega}} \frac{dz^0_{1,\omega}}{d\theta_{10}}$$

${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0,1$ and ${dz^0_{1,\omega}}/{d\theta_{10}} > 0$, but ${dz^1_{0,\omega}}/{d\theta_{10}} < 0$. The CEO span of control is nevertheless negative if $s_{0,\omega} > 1 - f_{0,\omega}$ and ${w_0}/{w_1} \geq {[1-f_{0,\omega}]}/{[1+f_{0,\omega}s_{1,\omega}/s_{0,\omega}]}$ (both sufficient, not necessary).

For $L_0 = L_1 = 2$:

$$\frac{d\sum_{j=0}^{1} n^{L_j}_{j,\omega}}{d\theta_{10}} = \sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{d\theta_{10}} \theta_{jj} e^{\lambda z^1_{j,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{\lambda z^1_{0,\omega}} \frac{dz^1_{0,\omega}}{d\theta_{10}} - n^0_{1,\omega}\theta_{11}\lambda e^{\lambda z^1_{1,\omega}} \frac{dz^1_{1,\omega}}{d\theta_{10}}$$

${dn^0_{j,\omega}}/{d\theta_{10}} < 0$ for $j = 0,1$ and ${dz^1_{1,\omega}}/{d\theta_{10}} > 0$, but ${dz^1_{0,\omega}}/{d\theta_{10}} < 0$. The CEO span of control is nevertheless negative if $s_{j,\omega} > 1 - f_{j,\omega}\, \forall j$ and ${w_0}/{w_1} \geq {[1-f_{0,\omega}]}/{[1+f_{0,\omega}s_{1,\omega}/s_{0,\omega}]}$ (both sufficient, not necessary).

To show: *The marginal benefit of CEO time $\varphi_{0,\omega}$ decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment's share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_{1,\omega}$.*

Follows from equation (C.3.68) due to $f_{1,\omega} \leq 1$ and ${d\bar{z}_{0,\omega}}/{d\theta_{10}} > 0$. □

C.3.6. Proposition 5: The optimal number of layers

Parameter values: $w_1 = w_0$, $\theta_{10} = \theta_{00}$, $\tau = 1$.

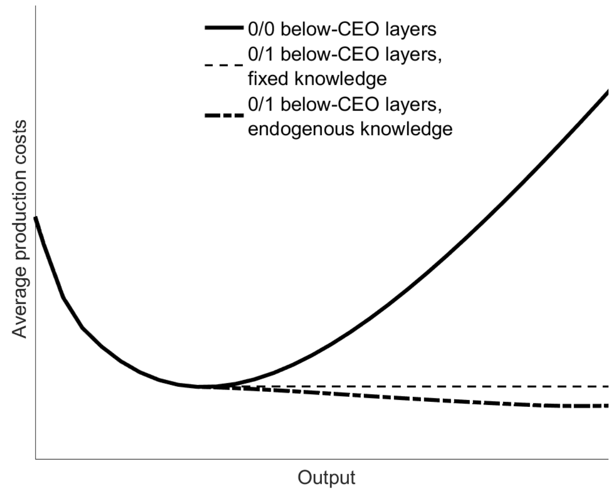
a) **To show:** *The average cost function of the (L_0, L_0)-organization is U-shaped in output and reaches a minimum at $\tilde{q}^{*(L_0,L_0)}$.*

   Follows from Proposition 1b): the cost function of the firm with the (L_0, L_0)-organization is equal to the cost function of a single-establishment firm with $n^0_{0,L_0} = \sum_{j=0}^{1} n^0_{j,(L_0,L_0)}$, because the firm chooses the same knowledge levels in the headquarters and the establishment by $\xi_{0,\omega} = \xi_{1,\omega}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ and $w_1 = w_0$, $\theta_{10} = \theta_{00}$.

b) **To show:** *The average cost of the (L_0, L_0+1)-organization and the (L_0+1, L_0)-organization coincide.*

   Follows from $w_1 = w_0$, $\theta_{10} = \theta_{00}$.

Figure C.1: Illustration: Proof of Proposition 5.



The figure illustrates part b) of the proof of Proposition 5. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{10} = \theta_{00} = .26$ (from Caliendo and Rossi-Hansberg 2012), $w_0 = w_1 = 1$. The solid line refers to an organization with $(0,0)$ below-CEO layers. The dashed lines show the average cost functions of organizations with $(0,1)$ below-CEO layers. The light dashed line refers to the organization with fixed knowledge levels, the bold dashed line to the organization with endogenous knowledge levels.

To show: *The average cost of the (L_0, L_0+1)-organization and the (L_0, L_0)-organization are equal at $\tilde{q}^{*(L_0,L_0)}$. The average cost function of the (L_0, L_0+1)-organization decreases with output $\tilde{q}$ for $\tilde{q}^{*(L_0+1,L_0+1)} > \tilde{q} > \tilde{q}^{*(L_0,L_0)}$.*

For simplicity and without loss of generality, we choose $L_0 = 0$.

The proof proceeds in two steps. First, we construct a $(0,1)$-organization at the minimum efficient scale of the$(0,0)$-organization $\tilde{q}^{*(0,0)}$. We fix knowledge levels and show that the $(0,1)$-organization produces $\tilde{q} \in [\tilde{q}^{*(0,0)}, \tilde{q}^{MAX}]$ with constant costs.

Second, we show that the average cost function of the organization with $(0,1)$ below-CEO layers and endogenous knowledge levels decreases with output for $\tilde{q} \in [\tilde{q}^{*(0,0)}, \tilde{q}^{*(1,1)})$. Figure C.1 illustrates the argument.

1. We construct a $(0,1)$-organization that has the same average cost as the $(0,0)$-organization at the minimum efficient scale $\tilde{q}^{*(0,0)}$.

   The knowledge levels of the $(0,0)$-organization coincide with the knowledge levels of a single establishment firm with no below-CEO layer (i.e., $L = 0$). Thus, at the minimum efficient scale $\tilde{q}^{*(0,0)}$,

$$\xi_{0,(0,0)} = \xi_{0,0} = AC_{0,0} \equiv AC_{0,0}^{MES} \tag{C.3.71}$$

$$\lambda z^0_{0,(0,0)} = \lambda z^0_{0,0} = \ln\left(\lambda \bar{z}_{0,0} + \frac{\lambda}{c}\right) + \ln\theta_{00} \equiv \lambda z^{0\,MES}_{0,0} \tag{C.3.72}$$

$$\begin{aligned}\lambda \bar{z}_{0,(0,0)} &= \lambda \bar{z}_{0,0} = \lambda z^0_{0,0} + \ln\left(\lambda z^0_{0,0} + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z^0_{0,0}}\right) - \ln\theta_{00} \\ &\equiv \lambda \bar{z}^{MES}_{0,0}\end{aligned} \tag{C.3.73}$$

$$\tilde{q}^*_{0,(0,0)} = \tilde{q}^*_{0,0} = \frac{1}{\theta_{00}} e^{\lambda z^0_{0,0}}(1 - e^{-\lambda \bar{z}_{0,0}})$$

Fix the knowledge levels of an ME firm with organization $\omega = (0,1)$ such that

$$z^0_{0,(0,1)} = z^{0\,MES}_{0,0}$$
$$\bar{z}_{0,(0,1)} = \bar{z}^{MES}_{0,0}$$
$$\xi_{1,(0,1)} = \xi_{0,0} \quad \Rightarrow \quad 1 + cz^0_{1,(0,1)} + \frac{c}{\lambda}\theta_{11}e^{-\lambda z^0_{1,(0,1)}} = 1 + cz^{0\,MES}_{0,0} \tag{C.3.74}$$
$$\varphi_{1,(0,1)} = \varphi_{0,0} \quad \Rightarrow \quad \theta_{11}e^{\lambda(z^1_{1,(0,1)} - z^0_{1,(0,1)})} = e^{\lambda z^{0\,MES}_{0,0}} \tag{C.3.75}$$
$$\text{with } z^1_{1,(0,1)} = \frac{1}{\lambda\theta_{11}}e^{\lambda z^0_{1,(0,1)}} - \frac{1}{c}$$

By construction, the average cost of the firm at $\tilde{q}^{*(0,0)}$ are $AC_{0,(0,1)} = AC^{MES}_{0,0}$.
The maximum producible quantity $\tilde{q}^{MAX}$ of the ME firm with organization $\omega = (0,1)$ and fixed knowledge levels is given by

$$\tilde{q}^{MAX} = \frac{1}{\theta_{00}}e^{\lambda z^1_{1,(0,1)}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}})$$

The ME firm produces both $\tilde{q}^{*(0,0)}$ and $\tilde{q}^{MAX}$ at the same costs, as at $\tilde{q}^{MAX}$,

$$\begin{aligned}
\xi_{1,(0,1)} =& \xi_{0,(0,1)} = \xi_{0,0} \quad \text{by construction} \\
AC_{0,(0,1)} =& w_0 \frac{1 + cz^0_{1,(0,1)} + \frac{c}{\lambda} + \theta_{00}e^{-\lambda z^1_{1,(0,1)}}(1 + c\bar{z}^{MES}_{0,0})}{1 - e^{-\lambda \bar{z}^{MES}_{0,0}}} \\
=& \xi_{0,(0,1)} - w_0 \frac{\frac{c}{\lambda}\theta_{11}e^{-\lambda z^0_{1,(0,1)}} - \theta_{00}e^{-\lambda z^1_{1,(0,1)}}(1 + c\bar{z}^{MES}_{0,0})}{1 - e^{-\lambda \bar{z}^{MES}_{0,0}}} \\
\overset{\varphi_{0,(0,1)} = \varphi_{1,(0,1)}}{=}& \xi_{0,(0,1)} - w_0 \frac{\theta_{00}\frac{c}{\lambda}e^{-\lambda z^0_{1,(0,1)}}e^{-\lambda z^{0\,MES}_{0,0}}\left(e^{\lambda z^0_{0,(0,1)}} - \theta_{00}\left(\frac{\lambda}{c} + \lambda\bar{z}^{MES}_{0,0}\right)\right)}{1 - e^{-\lambda \bar{z}^{MES}_{0,0}}} \\
=& \xi_{0,(0,1)} = AC^{MES}_{0,0} \quad \text{by (C.3.72)}
\end{aligned}$$

The ME firm produces output $\tilde{q}$ with $\tilde{q}^{MAX} \geq \tilde{q} \geq \tilde{q}^{*(0,0)}$ by allocating the share s of output to the headquarters and the share $1 - s$ to the establishment, where

$$s = \frac{\tilde{q} - \frac{1}{\theta_{00}}e^{\lambda z^1_{1,(0,1)}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}})}{\frac{1}{\theta_{00}}e^{\lambda z^{0\,MES}_{0,0}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}}) - \frac{1}{\theta_{00}}e^{\lambda z^1_{1,(0,1)}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}})}$$

The numerator and denominator are negative. The denominator is constant. $0 \leq s \leq 1$, because the numerator achieves its minimum at $\tilde{q} = \frac{1}{\theta_{00}}e^{\lambda z^{0\,MES}_{0,0}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}})$ $(s = 1)$, and its maximum at $\tilde{q} = \frac{1}{\theta_{00}}e^{\lambda z^1_{1,(0,1)}}(1 - e^{-\lambda \bar{z}^{MES}_{0,0}})$ $(s = 0)$.
That is, the average cost function of the ME firm with fixed knowledge levels is flat for $\tilde{q} \in \left[\tilde{q}^{*(0,0)}, \tilde{q}^{MAX}\right]$ (see the light dashed line in Figure C.1). □

2. We show that the average cost function of the organization with $(0,1)$ below-CEO layers and endogenous knowledge levels decreases with output for $\tilde{q} \in [\tilde{q}^{*(0,0)}, \tilde{q}^{*(1,1)})$. It is thus lower than the minimum average costs of the ME firm with $(0,0)$ below-CEO layers for $\tilde{q} > \tilde{q}^{*(0,0)}$, because it is lower than the average cost of an ME firm with organization $\omega = (0,1)$ and fixed knowledge levels.

The average cost of an ME firm with organization $\omega = (0,1)$ and optimal knowledge levels is lower than the average cost of the ME firm with organization ω but fixed knowledge levels because

$$C(\tilde{q}) \leq C(\tilde{q}, \bar{z}_{0,1}^{MES}, z_{0,1}^{0\,MES}, z_{1,(0,1)}^{0}(z_{0,1}^{0\,MES}), z_{1,(0,1)}^{1}(z_{0,1}^{0\,MES})).$$

The average cost function $AC_{0,(0,1)}(\tilde{q})$ decreases with output $\tilde{q}$ for $\tilde{q}^{*(1,1)} > \tilde{q} > \tilde{q}^{*(0,0)}$ by:

$$\frac{dAC_{0,\omega}(\tilde{q})}{d\tilde{q}} = \frac{1}{\tilde{q}}\left(\xi_{0,\omega} - AC_{0,\omega}\right) < 0 \text{ if } \xi_{0,\omega} < AC_{0,\omega}$$
$$\xi_{0,(0,1)} = \xi_{1,(0,1)} < AC_{0,(0,1)} \text{ if } \varphi_{0,(0,1)} = \varphi_{1,(0,1)} < w_0(1 + c\bar{z}_{0,(0,1)})$$

$\varphi_{j,(0,1)}$ is constant; $\bar{z}_{0,(0,1)}$ increases with $\tilde{q}$ by Proposition C.5. The maximum value of $AC_{0,(0,1)}(\tilde{q})$ is $AC_{0,0}^{MES}$. At $\tilde{q}^{*(0,0)}$, $\xi_{0,(0,1)} = AC_{0,0}$; $\xi_{0,(0,1)}$ decreases with $\tilde{q}$ for $\tilde{q}^{(0,1)\to(1,1)} > \tilde{q} > \tilde{q}^{*(0,0)}$. □

c) To show: *The average cost function of the (L_0+1, L_0+1)-organization intersects with the average cost function of the (L_0, L_0)-organization at the output $\tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)}$ between the minimum efficient scales, i.e., $\tilde{q}^{*(L_0+1,L_0+1)} > \tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)} > \tilde{q}^{*(L_0,L_0)}$. The average cost function of the (L_0, L_0+1)-organization intersects with the average cost function of the (L_0+1, L_0+1)-organization at a higher level of output $\tilde{q}^{(L_0,L_0+1)\to(L_0+1,L_0+1)} > \tilde{q}^{(L_0,L_0)\to(L_0+1,L_0+1)}$.*

We exploit the characteristics of the average cost function.

- $AC_{0,(0,1)} \leq AC_{0,0}^{MES}\, \forall\, \tilde{q}^{*(0,0)} \leq \tilde{q} \leq \tilde{q}^{MAX}$;
- $AC_{0,(0,0)}$ is increasing for $\tilde{q} > \tilde{q}^{*(0,0)}$;
- $AC_{0,(1,1)}$ is decreasing for $\tilde{q} \leq \tilde{q}^{*(1,1)}$, where $\tilde{q}^{MAX} \leq \tilde{q}^{*(1,1)}$;
- at $\tilde{q}^{*(0,0)}$, $AC_{0,(1,1)} > AC_{0,(0,0)}$.

In consequence, the increasing average costs function of the ME firm with $(0,0)$ below-CEO layers $AC_{0,(0,0)}$ intersects the decreasing average costs function of the ME firm with $(1,1)$ below CEO layers $AC_{0,(1,1)}$ at a lower output than the level at which the decreasing average cost function of the ME firm with $(0,1)$ below-CEO layers $AC_{0,(0,1)}$ intersects the average cost function $AC_{0,(1,1)}$. □

Corollary 1. *Suppose that wages are equal, $w_0 = w_1$, and that there are no transport costs or helping cost frictions, $\tau = 1, \theta_{00} = \theta_{10}$. When the multi-establishment firm adds a managerial layer at location j, knowledge levels $z_{j,\omega}^{\ell}$ at existing layers decrease discontinuously.*

Proof. For simplicity and without loss of generality, we choose $\omega = (0,0)$, and assume that the firm adds a layer at the establishment $(j = 1)$.

The corollary follows from equations C.3.74 and C.3.75, which imply that knowledge levels at the establishment decrease discontinuously. □

C.3.7. The optimal number of layers with transport frictions, wage and output differences

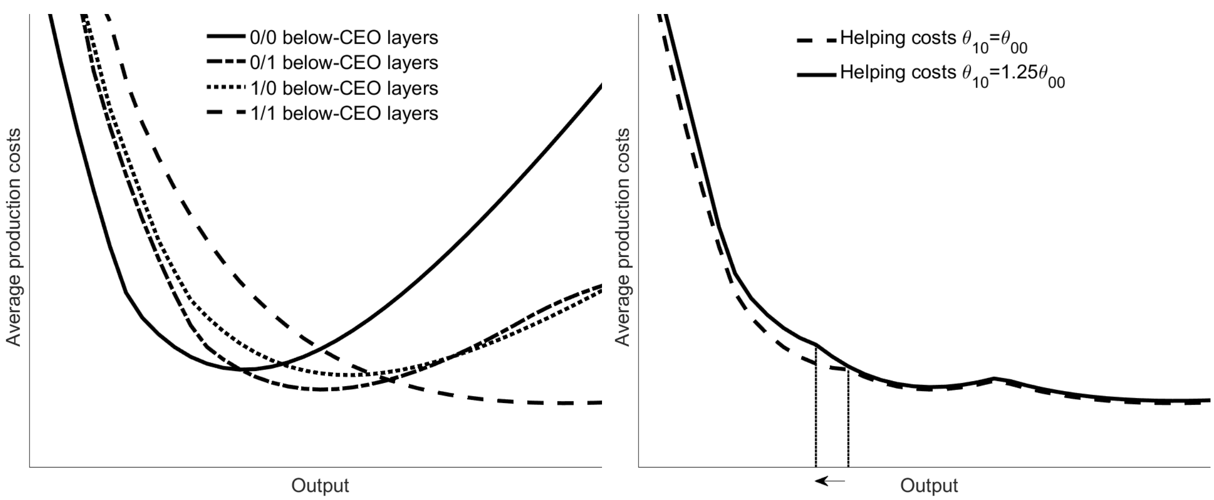
Overview. This section shows how wage and output differences between locations affect the optimal number and location of managerial layers. Most notably, higher wages or lower output at the establishment than at the headquarters can make it optimal to hire middle managers at the headquarters, but not the establishment.

Wage differences, $w_1 \neq w_0$, $\theta_{10} \geq \theta_{00}$, $\tau > 1$, $\tilde{q}_1 = \tilde{q}_0$. Lower wages at the establishment than the headquarters have the same effect as higher helping cost across space on the managerial organization. The firm first hires middle managers at the establishment, and then additionally at the headquarters as it grows, because middle managers at the establishment are cheaper. Higher helping costs reinforce this pattern (Figure C.2).

If wages are higher at the establishment than the headquarters, the firm adds a layer of middle managers first at the headquarters and then also at the establishment as it grows (Figure C.3a). Higher helping costs across space may outweigh the effect of higher wages at the establishment, so the firm adds a layer of middle managers only at the establishment (Figure C.3b). Still, the wage difference decreases the level of output at which middle managers at both units are optimal (Figure C.3c). Higher helping costs thus increase the number of managerial layers at the establishment and the headquarters.

Figure C.2: Lower wages at the establishment than at the headquarters

(a) $w_1 = 0.8w_0$, $\theta_{10} = \theta_{00}$ (b) $w_1 = 0.8w_0$, $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$

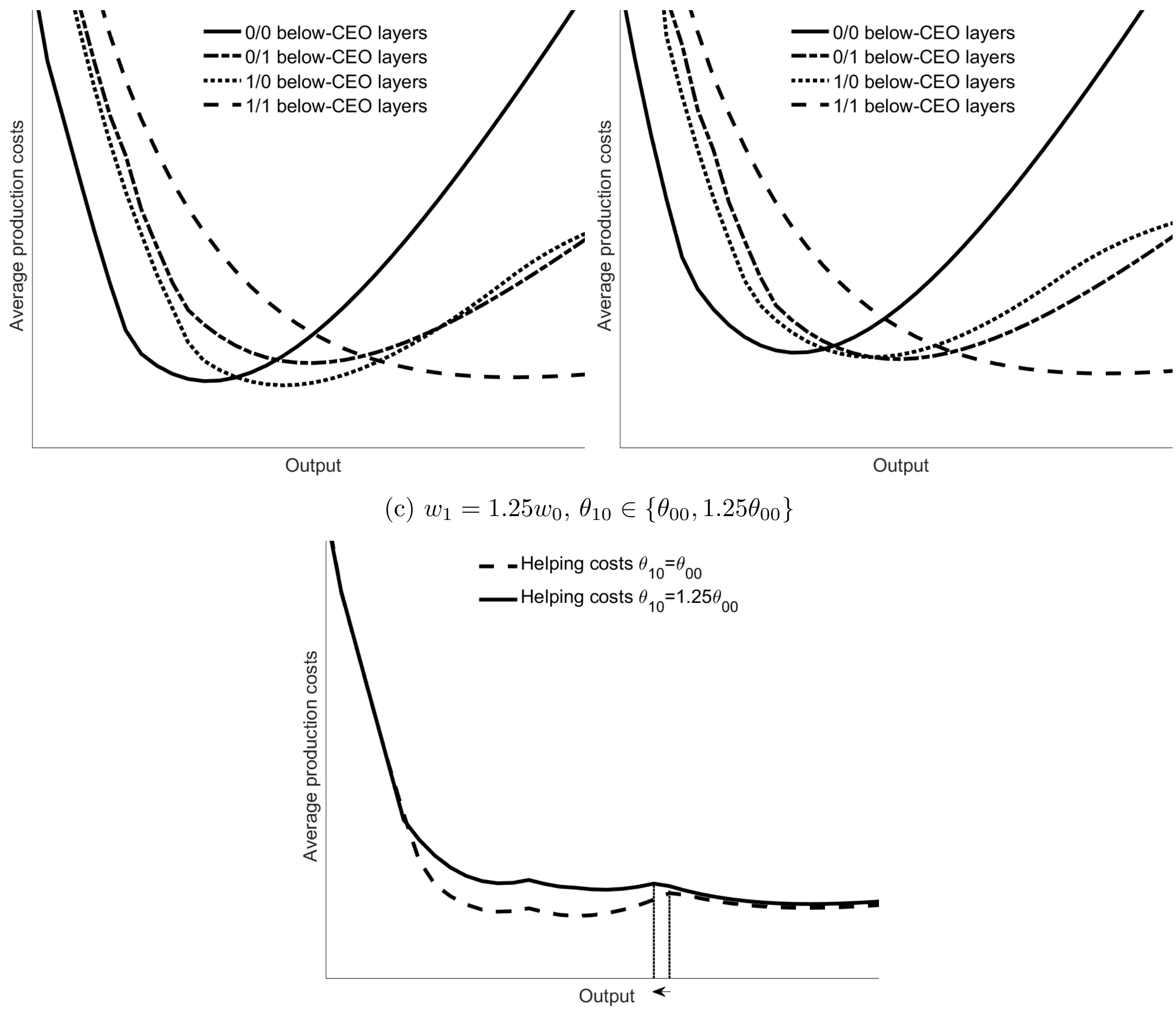


The figure plots the average cost functions of a ME firm for $w_1 = 0.8w_0$, $\tau = 1.5$, and $\tilde{q}_1 = \tilde{q}_0$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg 2012), $w_0 = 1$, $\tilde{q}_j \in [1, 63]$.
(a): $\theta_{10} = \theta_{00}$: Due to the lower wages at the establishment, the firm adds a layer of middle managers first only at the establishment and then also at the headquarters.
(b): $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$: Higher helping costs across space decrease the level of output at which the firm adds a layer of middle managers at the establishment.
Note: for the highest values of output in Figure (a), a layer of middle managers only at the headquarters has lower costs than a layer only at the establishment. The reason is that a firm with $(0,1)$-organization shifts total production to the establishment for high values of output and thus bears the transport costs, while it sticks to multi-establishment production with the $(1,0)$-organization. The costs of the $(0,1)$-organization are below those of the $(1,0)$-organization for even higher output.

Figure C.3: Higher wages at the establishment than at the headquarters

(a) $w_1 = 1.25w_0$, $\theta_{10} = \theta_{00}$ (b) $w_1 = 1.25w_0$, $\theta_{10} = 1.25\theta_{00}$

(c) $w_1 = 1.25w_0$, $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$



The figure plots the average cost functions of a ME firm for $w_1 = 1.25w_0$, $\tau = 1.5$, and $\tilde{q}_1 = \tilde{q}_0$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg 2012), $w_0 = 1$, $\tilde{q}_j \in [1, 63]$.

(a): $\theta_{10} = \theta_{00}$: Due to the lower wages at the headquarters, the firm adds a layer of middle managers first only at the headquarters and then also at the establishment.

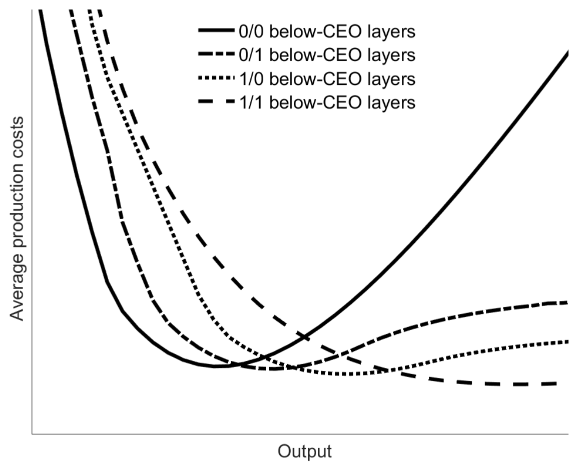
(b): $\theta_{10} = 1.25\theta_{00}$: The helping costs across space outweigh the higher wages at the establishment, so the firm adds a layer of middle managers only at the establishment for a range of output levels.

(c): $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$: Higher helping costs across space decrease the level of output at which the firm adds a layer of middle managers at the headquarters in addition to the establishment.

Note: for the highest values of output in the figure, adding a layer only at the establishment has lower average costs than adding it only at the headquarters. As in Figure C.2a, the reason is that the firm shifts total production to the headquarters with the $(1, 0)$-organization for high values of output and thus bears the transport costs.

Output differences, $\tilde{q}_1 \neq \tilde{q}_0$, $\tau > 1$, $\theta_{10} = \theta_{00}$, $w_1 = w_0$. We assume that headquarter output exceeds establishment output. As we assume equal location characteristics, the analogous results hold if establishment output exceeds headquarter output. If total output is close to the minimum efficient scale of the $(0,0)$-organization, hiring middle managers only at the establishment—the smaller unit—is optimal, because the quasi-fixed costs of the middle managers are low, but they release CEO time and thus decrease costs at the headquarters. For higher output levels, hiring middle managers only at the headquarters—the larger unit—is optimal because the firm saves the costs of middle managers at the establishment.

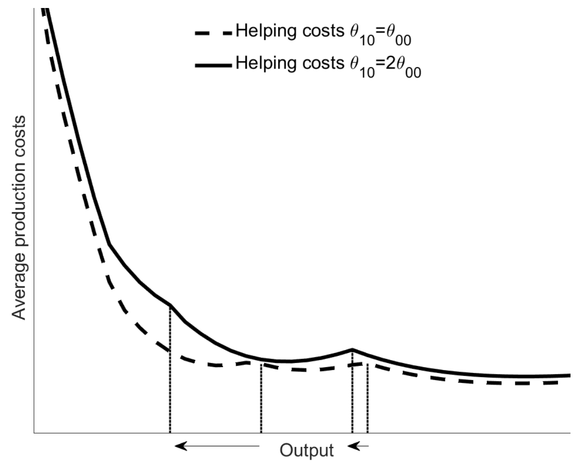
Figure C.4: Lower output at the establishment than at the headquarters



The figure plots the average cost functions of a ME firm for $\tilde{q}_1 = 0.5\tilde{q}_0$, $w_1 = w_0$, $\tau = 1.1$, and $\theta_{10} = \theta_{00}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg 2012), $w_0 = 1$, $\sum_{j=0}^{1} \tilde{q}_j \in [2, 126]$.

Helping cost and output differences, $\tau > 1$, $w_1 = w_0$. We assume that the helping costs across space increase and simultaneously establishment output decreases. Higher helping costs decrease the level of output at which adding a layer at the establishment and headquarters is optimal.

Figure C.5: Simultaneous change of helping costs and output



The figure plots the minimum average cost functions of a ME firm for $w_1 = w_0 = 1$, $\tau = 1.1$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg 2012), $\sum_{j=0}^{1} \tilde{q}_j \in [2, 126]$. Output on log scale. The dashed line assumes $\theta_{10} = \theta_{00}, \tilde{q}_1 = \tilde{q}_0$. The solid line assumes $\theta_{10} = 2\theta_{00}, \tilde{q}_1 = .66\tilde{q}_0$.

C.3.8. Comparative statics with respect to $\tilde{q}_j$, θ_{10} if $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $\tau \geq 1$

Proposition C.5. *Suppose the firm produces in the headquarters and the establishment. Suppose either that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other and that $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, or that there are no transport frictions $\tau = 1$, so $\xi_{0,\omega} = \xi_{1,\omega}$, but the headquarters and the establishment are not symmetric, i.e., $\theta_{10} \geq \theta_{00}$, and $w_1 < w_0$ or $L_1 \neq L_0$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$ increases with output $\tilde{q}_j$. Higher output $\tilde{q}_j$ increases the number of production workers $n^0_{k,\omega}$ and the share of CEO time $s_{k,\omega}$ at the location with the larger decrease of the marginal production costs and decreases their number and the share of CEO time at the other location, unless $L_0 > L_1$ and wages w_1 are too high.

b) The knowledge of the employees at all below-CEO layers $z^\ell_{k,\omega}$, $\ell \leq L_k$, the below-CEO managerial span of control $n^{\ell-1}_{k,\omega}/n^\ell_{k,\omega}$, $1 \leq \ell \leq L_k$, and the marginal benefit of CEO time $\varphi_{k,\omega}$, $k = 0, 1$, do not vary with output $\tilde{q}_j$.

c) The marginal production costs $\xi_{k,\omega}$, $k = 0, 1$, decrease with output $\tilde{q}_j$.

Under symmetry, i.e., $\tau = 1$, $\theta_{10} = \theta_{00}$, $w_1 = w_0$ and $L_1 = L_0$, output has the same effect on the choices of a multi-establishment firm as in Proposition 1.

Intuition.

- If $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, the multi-establishment firm effectively produces with two different technologies that have the same effective marginal production costs. It grows by recombining the technologies through reallocating CEO time and output.
- CEO knowledge increases with output, because knowledge and labor are complementary inputs. The number of production workers increases (decreases) at the location with the larger (smaller) decrease of the marginal costs, because the firm reallocates output to the location with the larger decrease of the marginal costs.
- The optimal combination of the knowledge levels of the employees at all below-CEO layers is uniquely given by the two conditions $\varphi_{0,\omega} = \varphi_{1,\omega}$ and $\xi_{0,\omega} = \xi_{1,\omega}$. Both conditions are independent of output. Correspondingly, neither knowledge levels nor the below-CEO span of control nor the marginal benefit of CEO time vary with output.

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$, $\tau \geq 1$. The second order conditions correspond to the ones in Appendix section C.3.4 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_j} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_j} = \tau\frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \frac{d\xi_{1,\omega}}{d\tilde{q}_j} = 0 \tag{C.3.76}$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_0} = 1 - \frac{dq_{0,\omega}}{d\tilde{q}_0} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_0} = 0 \tag{C.3.77}$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_1} = \tau - \frac{dq_{0,\omega}}{d\tilde{q}_1} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_1} = 0 \tag{C.3.78}$$

Proposition C.5, part a).

To show: *CEO knowledge $\bar{z}_{0,\omega}$ increases with output $\tilde{q}_0$, $\tilde{q}_1$ (or, if $\xi_{0,\omega} = \xi_{1,\omega}$, total output $\tilde{q}$).*

1. From equation (C.3.27):

$$\frac{dn^0_{k,\omega}}{d\tilde{q}_j} = \frac{1}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \left(\frac{dq_{k,\omega}}{d\tilde{q}_j} - n^0_{k,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} \right) \tag{C.3.79}$$

2. As will be shown below, $\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = 0$ and $\frac{dz^\ell_{k,\omega}}{d\tilde{q}_j} = 0$, $\ell = 0, ..., L_k$. From equation (C.3.26):

$$\frac{d\xi_{k,\omega}}{d\tilde{q}_j} = -\frac{\xi_{k,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} \tag{C.3.80}$$

3. Substituting equations (C.3.79) and (C.3.80) into equation (C.3.18) with $\tau \frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \frac{d\xi_{1,\omega}}{d\tilde{q}_j} = 0$ and $\mathbb{1}(j = 0)1 + \mathbb{1}(j = 1)\tau - \frac{dq_{0,\omega}}{d\tilde{q}_j} - \tau \frac{dq_{1,\omega}}{d\tilde{q}_j} = 0$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = \frac{\mathbb{1}(j = 0)1 + \mathbb{1}(j = 1)\tau}{(n^0_{0,\omega} + \tau n^0_{1,\omega})\lambda(1 + e^{-\lambda \bar{z}_{0,\omega}})} > 0$$

□

To show: *Higher output $\tilde{q}_j$ increases the number of production workers $n^0_{k,\omega}$ and the share of CEO time $s_{k,\omega}$ at the location with the larger decrease of the marginal production costs (i.e., the establishment) and decreases their number and the share of CEO time at the other location (i.e., the headquarters), unless $L_0 > L_1$ and wages w_1 are too high.*

From equations (C.3.19), (C.3.28) and (C.3.79):

$$\begin{aligned}
\frac{dq_{0,\omega}}{d\tilde{q}_0} &= \frac{\tau \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}}{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}} \\
\frac{dq_{1,\omega}}{d\tilde{q}_0} &= \frac{1}{\tau} \frac{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \tau \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0}}{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}} \\
\frac{dq_{0,\omega}}{d\tilde{q}_1} &= \frac{\tau \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1} - \tau \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}}{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}} \\
\frac{dq_{1,\omega}}{d\tilde{q}_1} &= \frac{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1}}{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}}
\end{aligned}$$

Substituting into equation (C.3.27) yields:

$$\begin{aligned}
\frac{dn^0_{0,\omega}}{d\tilde{q}_0} &= \frac{1}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0} \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}} (n^0_{0,\omega} + \tau n^0_{1,\omega}) - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}}{\tau \theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10} e^{-\lambda z^{L_1}_{1,\omega}}} \\
&< 0 \text{ by } \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0} < \frac{1}{\lambda e^{-\lambda \bar{z}_{0,\omega}} (n^0_{0,\omega} + \tau n^0_{1,\omega})}
\end{aligned}$$

$$
\begin{aligned}
\frac{dn^0_{1,\omega}}{d\tilde{q}_0} &= \frac{1}{\tau}\frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \tau\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega})}{\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}} \\
&> 0 \text{ by } \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_0} < \frac{1}{\lambda e^{-\lambda\bar{z}_{0,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega})} \\
\frac{dn^0_{0,\omega}}{d\tilde{q}_1} &= \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega}) - \tau\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}}{\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}} \\
&< 0 \text{ by } \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1} < \frac{\tau}{\lambda e^{-\lambda\bar{z}_{0,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega})} \\
\frac{dn^0_{1,\omega}}{d\tilde{q}_1} &= \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega})}{\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}} \\
&> 0 \text{ by } \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_1} < \frac{\tau}{\lambda e^{-\lambda\bar{z}_{0,\omega}}(n^0_{0,\omega}+\tau n^0_{1,\omega})}
\end{aligned}
$$

Note that $\frac{d\xi_{1,\omega}}{d\tilde{q}_j} < \frac{d\xi_{0,\omega}}{d\tilde{q}_j}$ by $\frac{d\xi_{1,\omega}}{d\tilde{q}_j} = \tau\frac{d\xi_{0,\omega}}{d\tilde{q}_j}$ and $\frac{d\xi_{k,\omega}}{d\tilde{q}_j} < 0$.
$\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$ if $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$:

- For $L_0 = L_1 \in \{0,1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$ if $w_0 > w_1$. For $w_0 = w_1$, multi-establishment production is only optimal if $\tau > 1$, so $\tau\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$.
- For $L_0 < L_1 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$ if $w_0 \geq w_1$.
- For $L_1 < L_0 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} > 0$ only if w_1 is sufficiently smaller than w_0. For $L_0 = 1$, $w_1 < w_0\theta_{00}e^{-\lambda z^0_{0,\omega}}$. For $L_0 = 2$, $w_1 < w_0 e^{\lambda(z^0_{1,\omega} - z^1_{0,\omega})}$ and $w_1 < w_0\theta_{00}e^{-\lambda z^1_{0,\omega}}$, respectively.

The results concerning the share of CEO time follow from equation (C.3.28). □

Proposition C.5, part b).

To show: *The knowledge of the employees at all below-CEO layers $z^{\ell}_{k,\omega}$, $\ell \leq L_k$, the below-CEO managerial span of control $n^{\ell-1}_{k,\omega}/n^{\ell}_{k,\omega}$, $1 \leq \ell \leq L_k$, and the marginal benefit of CEO time $\varphi_{k,\omega}$, $k = 0,1$, do not vary with output $\tilde{q}_j$.*
From equations (C.3.17) and (C.3.76):

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}_j} = \frac{d\varphi_{1,\omega}}{d\tilde{q}_j} \quad \text{and} \quad \tau\frac{d\xi_{0,\omega}}{d\tilde{q}_j} = \frac{d\xi_{1,\omega}}{d\tilde{q}_j}.$$

Substituting for $\frac{d\xi_{k,\omega}}{d\tilde{q}_j}$ from equation (C.3.26) implies, with $\tau\xi_{0,\omega} = \xi_{1,\omega}$:

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}_j}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} = \frac{d\varphi_{1,\omega}}{d\tilde{q}_j}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}$$

The equation holds if $\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = 0$ or $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} = \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}$, which contradicts $\varphi_{0,\omega} = \varphi_{1,\omega}$.

Equation (C.3.23) implies:

$$\begin{aligned}
&\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = \varphi_{k,\omega}\lambda\frac{dz^0_{k,\omega}}{d\tilde{q}_j} \text{ if } L_k = 0 \\
&\frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = \varphi_{k,\omega}\lambda\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} - w_k c\frac{\theta_{kk}}{\theta_{k0}}e^{\lambda(z^{L_k}_{k,\omega}-z^{L_k-1}_{k,\omega})}\frac{dz^{L_k-1}_{k,\omega}}{d\tilde{q}_j} \text{ if } L_k > 0 \\
&\Rightarrow \quad \frac{dz^\ell_{k,\omega}}{d\tilde{q}_j} = 0\,\forall k,\ell \text{ by } \frac{d\varphi_{k,\omega}}{d\tilde{q}_j} = 0\,\forall k
\end{aligned}$$

As the below-CEO managerial span of control is a function of knowledge, it is also constant. □

Proposition C.5, part c).

To show: *The marginal production cost $\xi_{k,\omega}$, $k = 0, 1$, decreases with output $\tilde{q}_j$ (or, if $\xi_{0,\omega} = \xi_{1,\omega}$, total output $\tilde{q}$).*
Follows from equation (C.3.80) and $\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} > 0$. □

Symmetry, i.e., $\tau = 1$, $\theta_{10} = \theta_{00}$, $w_1 = w_0$ and $L_1 = L_0$. The cost function coincides with the cost function of a single establishment firm, so Proposition 1 applies.

Case 2: $\xi_{0,\omega} = \tau\xi_{1,\omega}$. The second order conditions correspond to the ones in Appendix section C.3.4 with the following exceptions:

$$\begin{aligned}
\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_j} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_j} &= \frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \tau\frac{d\xi_{1,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_0} &= \tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_0} - \frac{dq_{1,\omega}}{d\tilde{q}_0} = 0 \\
\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_1} &= 1 - \tau\frac{dq_{0,\omega}}{d\tilde{q}_1} - \frac{dq_{1,\omega}}{d\tilde{q}_1} = 0
\end{aligned}$$

Proposition C.5, part a).

To show: *CEO knowledge $\bar{z}_{0,\omega}$ increases with output $\tilde{q}_0$, $\tilde{q}_1$.*
By an analogous argument to Case 1:

$$\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = \frac{\mathbb{1}(j=0)\tau + \mathbb{1}(j=1)1}{(\tau n^0_{0,\omega} + n^0_{1,\omega})\lambda(1 + e^{-\lambda\bar{z}_{0,\omega}})} > 0$$

□

To show: *Higher output $\tilde{q}_j$ increases the number of production workers $n^0_{k,\omega}$ and the share of CEO time $s_{k,\omega}$ at the location with the larger decrease of the marginal production costs and decreases their number and the share of CEO time at the other location, unless $L_0 > L_1$ and wages w_1 are too high.*
Follows from an analogous argument to Case 1. □

Proposition C.5, part b).

To show: *The knowledge of the employees at all below-CEO layers* $z^{\ell}_{k,\omega}$, $\ell \leq L_k$, *the below-CEO managerial span of control* $n^{\ell-1}_{k,\omega}/n^{\ell}_{k,\omega}$, $1 \leq \ell \leq L_k$, *and the marginal benefit of CEO time* $\varphi_{k,\omega}$, $k = 0, 1$, *do not vary with output* $\tilde{q}_j$.

Follows from an analogous argument to Case 1. ☐

Proposition C.5, part c).

To show: *The marginal production cost* $\xi_{k,\omega}$ *decreases with output* $\tilde{q}_j$.

Follows from an analogous argument to Case 1. ☐

Proposition C.6. *Suppose the firm produces in the headquarters and the establishment and that* $\theta_{10} > \theta_{00}$. *Suppose either that the firm incurs transport costs* $\tau > 1$ *to ship output from one location to the other and that* $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, *or that there are no transport frictions* $\tau = 1$, *so* $\xi_{0,\omega} = \xi_{1,\omega}$, *but the headquarters and the establishment are not symmetric, i.e.,* $w_1 < w_0$ *or* $L_1 \neq L_0$. *Given the organizational structure* ω,

a) *CEO knowledge* $\bar{z}_{0,\omega}$, *the marginal benefit of CEO time* $\varphi_{j,\omega}$, *the knowledge of the employees at all below-CEO layers* $z^{\ell}_{j,\omega}$, $\forall \ell < L_j$, *and the below-CEO managerial span of control* $n^{\ell-1}_{j,\omega}/n^{\ell}_{j,\omega}$, $1 \leq \ell \leq L_j$ *increase with the helping costs* θ_{10}.

b) *The number of production workers at the establishment* $n^0_{1,\omega}$ *decreases and the number of production workers at the headquarters* $n^0_{0,\omega}$ *increases with the helping costs* θ_{10}. *The total number of employees at all below-CEO layers* $\sum_{j=0}^{1} n^{\ell}_{j,\omega}$, $\forall \ell < L_j$ *decreases.*

c) *The marginal production cost* $\xi_{j,\omega}$ *increase with the helping costs* θ_{10}.

The comparative statics hold if $L_0 \leq L_1$, *or* $L_0 > L_1$ *and wages* w_1 *are sufficiently small.*

Intuition.

- If the marginal costs including transport costs are equal at the headquarters and the establishment, the firm reallocates output from the establishment to the headquarters in response to higher θ_{10}. In consequence, the number of production workers at the headquarters increases, as do their knowledge, the marginal benefit of CEO time and the marginal production costs.
- The intuition for the increase of CEO knowledge, the increase of the knowledge and the decrease of the number of employees at the establishment is analogous to Proposition 4.

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$**.** The second order conditions correspond to the ones in Appendix section C.3.5 with the following exceptions:

(C.3.81)
$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \tau\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$$

(C.3.82)
$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}} = -\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$$

Proposition C.6, part a).

To show: *CEO knowledge $\bar{z}_{0,\omega}$ increases with the helping costs θ_{10}.*

1. From equations (C.3.53) and (C.3.54), together with $\tau \frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$, $\tau\xi_{0,\omega} = \xi_{1,\omega}$, $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$ and $\varphi_{0,\omega} = \varphi_{1,\omega}$:

$$\text{(C.3.83)} \qquad \frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{\theta_{00}\varphi_{0,\omega} - \xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}}(\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}})}{(1 - e^{-\lambda\bar{z}_{0,\omega}})(\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}})}$$

2. Substituting equation (C.3.83) into equation (C.3.53) yields:

$$\text{(C.3.84)} \qquad \frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}e^{\lambda z_{0,\omega}^{L_0}}}{\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}}}$$

3. From equation (C.3.55):

$$\text{(C.3.85)} \qquad \frac{dn_{j,\omega}^0}{d\theta_{10}} = \frac{\frac{dq_{j,\omega}}{d\theta_{10}} - n_{j,\omega}^0\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}}{1 - e^{-\lambda\bar{z}_{0,\omega}}}$$

4. Substituting equation (C.3.85) together with $\tau\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$ and $-\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$ into equation (C.3.44) yields:

$$\text{(C.3.86)} \qquad \frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{d\bar{z}_{0,\omega}}{d\theta_{10}}\frac{\xi_{0,\omega}\lambda}{1 - e^{-\lambda\bar{z}_{0,\omega}}}$$

5. Combining equations (C.3.83) and (C.3.86) yields:

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}\theta_{00}}{\lambda\xi_{0,\omega}(1 + e^{-\lambda\bar{z}_{0,\omega}})(\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}})}$$

   $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ if $\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}} > 0$. This expression holds for $L_0 \leq L_1$ and for $L_0 > L_1$ unless wages w_1 are too high, see proof of Proposition C.5, part a. □

To show: *The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with the helping costs θ_{10}.* Follows from equation (C.3.84). The expression is positive if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$. □

To show: *The knowledge of the employees at all below-CEO layers $z_{j,\omega}^{\ell}$, $\forall \ell \leq L_j$ and the below-CEO managerial span of control $n_{j,\omega}^{\ell-1}/n_{j,\omega}^{\ell}$, $1 \leq \ell \leq L_j$ increase with the helping costs θ_{10}.*

Follows from $\frac{d\varphi_{j,\omega}}{d\theta_{10}} > 0$ and equations (C.3.49, C.3.50,) which imply:

$$\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} \overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$$\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} - \frac{dz_{0,\omega}^{L_0-1}}{d\theta_{10}} \overset{L_0>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} - \frac{dz_{1,\omega}^{L_1-1}}{d\theta_{10}} \overset{L_1>0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

Knowledge at lower layers depends on knowledge at higher layers as in the single-establishment firm (see section C.2.3). The below-CEO managerial span of control increases, because it is a positive function of knowledge. □

Proposition C.6, part b).

To show: *The number of production workers at the establishment $n_{1,\omega}^0$ decreases and the number of production workers at the headquarters $n_{0,\omega}^0$ increases with the helping costs θ_{10}.*

Substituting equation (C.3.55) into equation (C.3.82) yields:

$$\frac{dn_{1,\omega}^0}{d\theta_{10}} = -\frac{1}{\tau}\frac{dn_{0,\omega}^0}{d\theta_{10}} - \frac{1}{\tau}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}(n_{0,\omega}^0 + \tau n_{1,\omega}^0)\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}$$

Substituting this expression into equation (C.3.45) together with equations (C.3.56, C.3.57) yields:

$$\frac{dn_{0,\omega}^0}{d\theta_{10}} = \frac{e^{\lambda z_{0,\omega}^{L_0}} e^{\lambda z_{1,\omega}^{L_1}}}{\tau\theta_{00}e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10}e^{\lambda z_{0,\omega}^{L_0}}} \times \left(\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}(n_{0,\omega}^0 + \tau n_{1,\omega}^0)\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \tau\sum_{j=0}^{1}\lambda s_{j,\omega}\frac{dz_{j,\omega}^{L_j}}{d\theta_{10}} - \tau n_{1,\omega}^0 e^{-\lambda z_{1,\omega}^{L_1}} \right) > 0$$

by $\lambda s_{1,\omega}\frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} > n_{1,\omega}^0 e^{-\lambda z_{1,\omega}^{L_1}}$ as $\frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} > \frac{1}{\lambda\theta_{10}}$. □

To show: *The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n_{j,\omega}^{\ell}$, $\forall \leq L_j$ decreases with the helping costs θ_{10}.*
$\ell = 0$: Follows from equation (C.3.55), , with $-\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$:

□
$$\sum_{j=0}^{1}\frac{dn_{j,\omega}^0}{d\theta_{10}} = -\sum_{j=0}^{1}\frac{n_{j,\omega}^0\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} < 0 \quad \text{as} \quad \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$$

$\ell > 0$: Follows from $\sum_{j=0}^{1}\frac{dn_{j,\omega}^0}{d\theta_{10}} < 0$ and $\frac{dz_{j,\omega}^{\ell}}{d\theta_{10}} > 0$. □

Proposition C.6, part c).

To show: The marginal production cost $\xi_{j,\omega}$ increase with the helping costs θ_{10}.
Follows from $\frac{d\xi_{1,\omega}}{d\theta_{10}} = \tau\frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{d\bar{z}_{0,\omega}}{d\theta_{10}}\frac{\xi_{0,\omega}\lambda}{1-e^{-\lambda\bar{z}_{0,\omega}}} > 0$ (equation C.3.86) if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$. $\frac{d\xi_{1,\omega}}{d\theta_{10}} = \tau\frac{d\xi_{0,\omega}}{d\theta_{10}}$ implies that $\frac{d\xi_{1,\omega}}{d\theta_{10}} > \frac{d\xi_{0,\omega}}{d\theta_{10}}$. □

Case 2: $\xi_{0,\omega} = \tau\xi_{1,\omega}$. The second order conditions correspond to the ones in Appendix section C.3.5 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} - \tau\frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\theta_{10}} = -\tau\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$$

Results follow from derivations analogous to those for Case 1. □

C.4. The optimal output

Proposition 6. The profit maximization problem and the first order conditions are given by:

$$\max_{\tilde{q}_0,\tilde{q}_1\geq 0} \pi_i = \sum_{j=0}^{1} p_j(\tilde{q}_j)\tilde{q}_j - C(\tilde{q}_0,\tilde{q}_1)$$
$$\frac{\partial\pi_i}{\partial\tilde{q}_j} = \frac{\partial p_j}{\partial\tilde{q}_j}\tilde{q}_j + p_j(\tilde{q}_j) - \xi_{j,\omega} = 0$$

$\tau\xi_{j,\omega} \neq \xi_{k,\omega}$. We define $\hat{q}_0 \equiv -\tilde{q}_0$. From Proposition 4:

$$\frac{\partial^2\pi_i}{\partial\hat{q}_0\partial\theta_{10}} = \frac{\partial\xi_{0,\omega}}{\partial\theta_{10}} < 0$$
$$\frac{\partial^2\pi_i}{\partial\tilde{q}_1\partial\theta_{10}} = -\frac{\partial\xi_{1,\omega}}{\partial\theta_{10}} < 0$$

By monotone comparative statics, $\hat{q}_0$ and $\tilde{q}_1$ decrease with the helping costs θ_{10} if

$$\frac{\partial^2\pi_i}{\partial\hat{q}_0\partial\tilde{q}_1} = \frac{\partial\xi_{0,\omega}}{\partial\tilde{q}_1} > 0$$

This holds for sufficiently high output $\tilde{q}_j$. In result, $\tilde{q}_0$ increases and $\tilde{q}_1$ decreases with the helping costs θ_{10}. □

$\tau\xi_{j,\omega} = \xi_{k,\omega}$. From Proposition C.6:

$$\frac{\partial^2\pi_i}{\partial\tilde{q}_j\partial\theta_{10}} = -\frac{\partial\xi_{j,\omega}}{\partial\theta_{10}} < 0$$

From Proposition C.5:

$$\frac{\partial^2\pi_i}{\partial\tilde{q}_0\partial\tilde{q}_1} = -\frac{\partial\xi_{0,\omega}}{\partial\tilde{q}_1} > 0$$

By monotone comparative statics, both $\tilde{q}_0$ and $\tilde{q}_1$ decrease with the helping costs θ_{10}. □

D. Reorganization due to high-speed railway routes

D.1. Model predictions

Table D.1: Total effect of lower helping costs θ_{10} given the organizational structure ω

Variable	Direct effect *Prop. 4*	Effect on $\tilde{q}_0$ *Prop. 6*	$\times$ Effect of $\tilde{q}_0$ *Prop. 3*	Effect on $\tilde{q}_1$ *Prop. 6*	$\times$ Effect of $\tilde{q}_1$ *Prop. 3*	Total effect
Firm-level variables						
CEO knowledge $\bar{z}_{0,\omega}$	-	-	+	+	+	+/-
CEO time at HQ $s_{0,\omega}$	(-)	-	+	+	-	-
CEO time at est. $s_{1,\omega}$	(+)	-	-	+	+	+
CEO span $\sum_{j=0}^{1} n_{j,\omega}^{L_j}$	+	-	+/-	+	+/-	+/-
Production at HQ $q_{0,\omega}$	0	-	+	+	0	-
Production at est. $q_{1,\omega}$	0	-	0	+	+	+
Establishment/Headquarter-level variables						
# workers at HQ $n_{0,\omega}^{0}$	+	-	+	+	-	+/-
# workers at est. $n_{1,\omega}^{0}$	+	-	-	+	+	+
Knowledge at HQ $z_{0,\omega}^{\ell}$	+	-	+	+	+	+/-
Knowledge at est. $z_{1,\omega}^{\ell}$	-	-	+	+	+	+/-
Span at HQ $n_{0,\omega}^{\ell}/n_{0,\omega}^{\ell+1}$	+	-	+	+	+	+/-
Span at est. $n_{1,\omega}^{\ell}/n_{1,\omega}^{\ell+1}$	-	-	+	+	+	+/-

The table displays the direct and indirect effects of lower helping costs on the endogenous variables given the organizational structure ω. $+\,(+)$ denotes (weakly) positive effects, $-\,(-)$ denotes (weakly) negative effects, $+/-$ denotes ambiguous effects, and 0 denotes no effect.

In supplementary analyses, we calibrate our model to moments of the data on firms with one establishment. We simulate the effect of a reduction of the helping costs across space. The simulation delivers three main insights:

1. While the sign of the direct and total effect of a reduction of the helping costs is similar across organizational structures, the size of the effects depends on the organizational structure.

2. The indirect effect of lower helping costs through endogenous changes of output can be quantitatively larger than their direct effect.

3. The indirect effect can be strong enough to outweigh the direct effect in case the signs of the two effects differ.

D.2. Background information on high-speed railway routes

Travel time data. Our data comprise 115 train stations that are connected to the long-distance network in at least one of the years 2000, 2004 and 2008. To ensure that temporary construction works do not affect travel times, Deutsche Bahn AG computed the travel times for three different Wednesdays in March, June and November. We assume that passengers leave between 5 am and 7 pm. Travel times may change for several reasons, such as adjustments to time tables, construction works, or new changeover connections. To allow us to disentangle lower travel times due to the new routes and other reasons, the data contain an indicator for station pairs where more than 50 percent of passengers used one of the new routes in 2008.

Effect of high-speed routes on travel times. Table D.2 displays the change of the minimum travel times when we use the indicator to distinguish county pairs where passengers are more or less likely to use the new high-speed routes. As the table shows, minimum travel times change very little between counties where less than half of all passengers use the new high-speed routes. Travel times decrease substantially between counties where passengers frequently use the new high speed routes, although not all county pairs benefit to the same extent from the new routes due to the complexity of the German long-distance railway network. To illustrate this with a specific example, route 3 between Ingolstadt and Nuremberg substantially decreased the travel time between Munich and Nuremberg as well as cities north of Nuremberg such as Würzburg and Leipzig. The route did not decrease travel times as much between Munich and Frankfurt, even though Frankfurt is located north of Nuremberg, because of an alternative connection between the cities via Stuttgart.

Table D.2: Reduction of travel times in minutes

	N	HSR	Mean	SD	p25	p50	p75
2000-2004	2,876	0	-2.5	17.6	-7	0	4
	987	1	-32.8	43.7	-59	-26	1
2004-2008	2,913	0	-0.3	15.1	-3	0	4
	1,023	1	-11.9	25.1	-23	-6	1

The table displays summary statistics on the reduction of travel time between 2000 and 2004 and 2004 and 2008 separately for connections where at least 50% of passengers use the new high-speed routes (HSR) and other routes. The level of observation is a pair of headquarter county and county.

Construction of high-speed routes. Deutsche Bahn AG (strictly speaking, its predecessor Bundesbahn) started operating high-speed trains with a speed of up to 250km/h on selected routes in June 1991. Trains are commonly known as ICE ("InterCity Express") trains. Their speed depends on the type of route:

- 250-300km/h on newly constructed routes ("*Neubaustrecken*"),
- up to 200km/h on upgraded routes ("*Ausbaustrecken*"),
- up to 160km/h on normal long-distance routes.

During the sample period, Deutsche Bahn AG opened two newly constructed routes between Frankfurt and Cologne and Ingolstadt and Nuremberg, and upgraded the route between Hamburg and Berlin. The maximum speed on this route is 230km/h, so it is commonly considered a high-speed route (Jänsch 2006).

Route 1: Frankfurt-Cologne. Construction of route 1 started in 1995 and was completed after six years. The route almost halved the travel time between Frankfurt and Cologne from 135 to 76 minutes. Trains started operating as shuttles between Frankfurt and Cologne in August 2002. The route became part of the railway network in December 2002 (Brux 2003). The route runs in parallel to the autobahn A3. More than 100 lawsuits against the results of the plan approval procedures led to numerous changes during the construction period (Kaniut and Form 2002). In addition, execution errors and quality defects resulted in delays. For example, parts of a newly constructed bridge had to be demolished due to defective concrete (Belter 2004).

Route 2: Hamburg-Berlin. The first railway route between Hamburg and Berlin was constructed in 1844-1846. After reunification, its upgrade became part of the infrastructure project "*Deutsche Einheit*" (German unity). A first upgrade was completed in 1997 and reduced the travel time to 134 minutes at a maximum speed of 160km/h. In 2000, the federal government decided to provide an additional upgrade to the existing route instead of building a magnetic levitation train between Hamburg and Berlin. Train service on the upgraded route started in December 2004. The route reduced the travel time to 90 minutes (Feldwisch et al. 2004).

Route 3: Ingolstadt-Nuremberg. First construction works for route 3 started in 1994; extensive construction started in 1998 (Weigelt 2003). The route reduced the travel time between Ingolstadt and Nuremberg from 66 to 30 minutes from May 2006 onwards (Brux 2006). After being connected to the upgraded route between Ingolstadt and Munich, the route also reduced the travel time between Munich and Nuremberg from December 2006 (Feldwisch and Schülke 2006). The route runs in parallel to the autobahn A9 (Weigelt 2003). About a third of the route consists of tunnels. Unexpected and partly unforeseeable difficulties (in particular due to karst) led to delays during construction (Brux 2006; Wegener 2003).

Freight traffic. In general, passenger and freight trains use the same tracks in the German railway network (so-called "*Mischverkehr*" or mixed traffic). For safety reasons, passenger trains run during the day and freight trains mostly during the night (Jänsch 2010).

However, the routes between Frankfurt and Cologne and Nuremberg and Ingolstadt are exceptions to this rule. The route between Frankfurt and Cologne was designed exclusively for passenger traffic from the very beginning (Jänsch 2015). The route between Nuremberg and Ingolstadt was designed for mixed traffic, but was ultimately not licensed for freight traffic. The reason is that there are tunnels on the route. High-speed passenger and freight trains must not use a two-track tunnel in different directions at the same time for safety reasons, and there is no technology to preclude this (Feldwisch and Schülke 2006). According to official records from the German parliament, not a single freight train used the route during our sample period (Deutscher Bundestag 2010). It is thus unlikely that the new high-speed routes reduced trade costs between locations.

The route between Hamburg and Berlin is the only route used by freight trains. Our results are similar if we exclude this route (see Table D.17).

Effects on air traffic. The attractiveness of high-speed railway routes is reflected in their effects on air traffic. For example, a regular plane service between Cologne Bonn Airport and Frankfurt Airport was discontinued in 2007. The carrier Lufthansa cited the new high-speed railway route as the main reason for lower demand (Eurailpress.de 2007). The number of flights between Cologne Bonn Airport and Nuremberg Airport has also dropped substantially (Deutscher Bundestag 2007).

D.3. Descriptive statistics

Table D.3: Descriptive statistics, 2000-2010 panel

Descriptive statistics	Unit	N	Mean	SD	p25	p50	p75
Log # non-managerial employees	D-E	47,732	2.1	1.5	1.1	1.9	3.0
	HQ	13,393	3.7	1.4	2.8	3.7	4.6
	ND-E	45,508	2.1	1.5	1.1	1.9	3.0
Average log non-managerial wages	D-E	47,732	4.5	0.4	4.3	4.5	4.7
	HQ	13,393	4.6	0.3	4.4	4.6	4.8
	ND-E	45,508	4.4	0.3	4.3	4.5	4.7
# managerial layers	D-E	47,732	0.8	1.0	0	1	1
	HQ	13,393	1.9	1.1	1	2	3
	ND-E	45,508	0.7	0.9	0	0	1
Managerial share	D-E	47,732	6.2	15.2	0	0	3.8
(%, Blossfeld)	HQ	13,393	9.6	14.0	0	4.6	12.1
	ND-E	45,508	5.4	14.3	0	0	2.5
Establishment treated	D-E	47,732	0.11	0.32	0	0	0
Firm treated	HQ	13,393	0.20	0.40	0	0	0
	ND-E	45,508	0.38	0.48	0	0	1

Summary statistics for Table VII. *D-E:* directly affected establishment, *HQ:* headquarters, *ND-E:* non-directly affected establishment.

D.4. Evidence supporting the identification strategy

Table D.4: Robustness to alternative control groups and strategic location of establishments (Figure VIII), 2000-2010 panel

Directly affected establishments								
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alternative control groups								
	Firms with one establishment				Sample w/o non-directly affected est.			
Est. treated	0.146***	−0.007	−0.004	0.231	0.043*	0.011*	−0.003	0.281
	(0.040)	(0.013)	(0.040)	(0.802)	(0.020)	(0.004)	(0.014)	(0.250)
# observations	6,049	6,049	6,049	6,049	29,661	29,661	29,661	29,661
# est.	809	809	809	809	3,336	3,336	3,336	3,336
R-squared	0.930	0.933	0.892	0.869	0.921	0.921	0.873	0.863
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Entry before 2000								
	All firms				Firms with $\geq$ 2 establishments			
Est. treated	0.073**	−0.001	−0.004	0.161	0.068**	0.000	0.006	0.139
	(0.022)	(0.004)	(0.017)	(0.298)	(0.023)	(0.005)	(0.020)	(0.290)
# observations	41,255	41,255	41,255	41,255	34,807	34,807	34,807	34,807
# est.	4,666	4,666	4,666	4,666	3,986	3,986	3,986	3,986
R-squared	0.911	0.933	0.882	0.899	0.911	0.933	0.887	0.912
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Entry before 1995								
	All firms				Firms with $\geq$ 2 establishments			
Est. treated	0.085**	−0.001	−0.016	−0.188	0.082**	−0.002	−0.012	−0.204
	(0.026)	(0.006)	(0.021)	(0.302)	(0.031)	(0.007)	(0.027)	(0.297)
# observations	23,900	23,900	23,900	23,900	19,947	19,947	19,947	19,947
# est.	2,677	2,677	2,677	2,677	2,254	2,254	2,254	2,254
R-squared	0.918	0.926	0.885	0.904	0.918	0.926	0.890	0.917
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by county in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile and year in the top panel and by size quartile, county and year in the middle and bottom panels. All variables are winsorized at the first and 99th percentiles.

Table D.5: The effect of opening of high-speed railway routes, event-study estimates (Figure IX), 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Year -4/-3	−0.026	0.008	−0.003	−0.211	−0.003	0.005	0.012	−0.063
	(0.023)	(0.009)	(0.023)	(0.195)	(0.023)	(0.012)	(0.024)	(0.170)
Year -2/-1	−0.024^{+}	0.002	−0.001	−0.012	−0.008	0.000	−0.006	−0.056
	(0.014)	(0.004)	(0.011)	(0.139)	(0.014)	(0.003)	(0.012)	(0.177)
Year 1/2	0.059***	−0.001	0.000	−0.153	0.072***	0.001	0.005	−0.110
	(0.014)	(0.003)	(0.011)	(0.110)	(0.014)	(0.003)	(0.013)	(0.120)
Year 3/4	0.061**	0.006	0.006	0.287	0.084***	0.006	0.014	0.213
	(0.021)	(0.006)	(0.016)	(0.300)	(0.023)	(0.007)	(0.018)	(0.275)
Year 5/6	0.030	0.012	−0.010	−0.090	0.041	0.012	−0.003	0.017
	(0.033)	(0.008)	(0.031)	(0.401)	(0.039)	(0.010)	(0.035)	(0.388)
# observations	43,815	43,815	43,815	43,815	36,904	36,904	36,904	36,904
# est.	4,530	4,530	4,530	4,530	3,840	3,840	3,840	3,840
R-squared	0.905	0.937	0.877	0.902	0.904	0.937	0.883	0.914
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Year -2/-1	−0.011	−0.002	−0.012	−0.235	−0.026	−0.008*	−0.014	−0.308
	(0.013)	(0.002)	(0.014)	(0.230)	(0.022)	(0.004)	(0.027)	(0.396)
Year 1/2	−0.010	−0.002	0.028*	−0.124	0.005	0.005	0.047**	0.348
	(0.013)	(0.003)	(0.013)	(0.167)	(0.020)	(0.004)	(0.015)	(0.242)
Year 3/4	−0.037^{+}	0.000	−0.048^{+}	−0.348	−0.009	0.011^{+}	0.006	0.119
	(0.019)	(0.005)	(0.026)	(0.260)	(0.047)	(0.005)	(0.051)	(0.417)
# observations	12,187	12,187	12,187	12,187	6,281	6,281	6,281	6,281
# HQ	1,296	1,296	1,296	1,296	687	687	687	687
R-squared	0.955	0.953	0.879	0.924	0.955	0.950	0.879	0.931
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Year -2/-1					−0.004	−0.001	−0.002	−0.030
					(0.010)	(0.003)	(0.009)	(0.149)
Year 1/2					−0.017	0.003	−0.002	0.134
					(0.010)	(0.003)	(0.009)	(0.144)
Year 3/4					−0.059*	0.028***	0.001	0.653^{+}
					(0.028)	(0.008)	(0.023)	(0.380)
# observations					47,245	47,245	47,245	47,245
# est.					4,359	4,359	4,359	4,359
R-squared					0.923	0.930	0.897	0.895
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (headquarter) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Coefficients and standard errors from regressions similar to equations (26)-(28). *Dependent variables:* see Table VII. *Independent variables:* indicator variables for lower travel times to the HQ (top panel) and lower travel times between HQ and at least one establishment (middle and bottom panels) interacted with biannual fixed effects. Treated and control units are matched by size quartile, (headquarter) county and year (top and middle panel) and size quartile and year (bottom panel). All variables are winsorized at the first and 99th percentiles.

The results in column 6 of Table D.5 suggest that wages in treated headquarters grow more slowly than wages in control headquarters before the opening of the routes. This effect is not robust, however, as the following table with annual interaction terms shows:

Event-study estimates, HQ of firms with ≥ 2 est.s, wages, 2000-2010 panel

	Year -2	Year -1	Year 1	Year 2	Year 3	Year 4
wages	-0.018	-0.003	0.003	0.008	0.012*	0.010
	(0.011)	(0.002)	(0.004)	(0.007)	(0.005)	(0.007)

Results for column 6, headquarters, in Table D.5 with annual interaction terms. # observations: 6,281. # HQ: 687.

D.5. Robustness checks

D.5.1. Weighting, statistical inference, and choice of strata

No weights. Tables D.6 and D.7 document that results in Table VII and Figure VIII are similar if we do not use the weights recommended by Iacus et al. (2012) to estimate the average treatment effect on the treated.

Table D.6: Table VII without weights, 2000–2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em. (1)	wages (2)	# lay. (3)	mg.sh. (4)	# em. (5)	wages (6)	# lay. (7)	mg.sh. (8)
Directly affected establishment								
Est. treated	0.074***	0.001	0.001	0.057	0.068***	0.003	0.008	0.091
	(0.017)	(0.004)	(0.015)	(0.232)	(0.019)	(0.005)	(0.017)	(0.260)
# observations	47,732	47,732	47,732	47,732	40,143	40,143	40,143	40,143
# est.	5,609	5,609	5,609	5,609	4,791	4,791	4,791	4,791
R-squared	0.906	0.924	0.877	0.884	0.908	0.926	0.881	0.895
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.021	0.007	0.028	0.157	0.030	0.015*	0.057*	0.618
	(0.017)	(0.005)	(0.022)	(0.323)	(0.034)	(0.006)	(0.026)	(0.464)
# observations	13,393	13,393	13,393	13,393	6,261	6,261	6,261	6,261
# HQ	1,469	1,469	1,469	1,469	683	683	683	683
R-squared	0.952	0.947	0.878	0.902	0.954	0.950	0.879	0.911
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.012	0.018***	0.022	0.516*
					(0.021)	(0.004)	(0.017)	(0.203)
# observations					45,508	45,508	45,508	45,508
# est.					5,508	5,508	5,508	5,508
R-squared					0.916	0.926	0.890	0.888
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. Observations are not weighted. All variables are winsorized at the first and 99th percentiles.

Table D.7: Table D.4 without weights, 2000-2010 panel

Directly affected establishments								
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alternative control groups								
	Firms with one establishment				Sample w/o non-directly affected est.			
Est. treated	0.146***	−0.008	−0.004	0.187	0.040*	0.011**	−0.004	0.269
	(0.040)	(0.013)	(0.041)	(0.766)	(0.020)	(0.004)	(0.014)	(0.249)
# observations	6,049	6,049	6,049	6,049	29,661	29,661	29,661	29,661
# est.	809	809	809	809	3,336	3,336	3,336	3,336
R-squared	0.930	0.933	0.890	0.869	0.920	0.920	0.872	0.861
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Entry before 2000								
	All firms				Firms with ≥ 2 establishments			
Est. treated	0.061**	−0.002	−0.007	0.160	0.053*	0.000	0.001	0.164
	(0.019)	(0.925)	(0.881)	(0.254)	(0.021)	(0.005)	(0.018)	(0.286)
# observations	41,255	41,255	41,255	41,255	34,807	34,807	34,807	34,807
# est.	4,666	4,666	4,666	4,666	3,986	3,986	3,986	3,986
R-squared	0.914	0.925	0.881	0.891	0.916	0.926	0.885	0.902
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Entry before 1995								
	All firms				Firms with ≥ 2 establishments			
Est. treated	0.072**	−0.002	−0.019	−0.052	0.072**	−0.002	−0.011	−0.064
	(0.025)	(0.006)	(0.021)	(0.293)	(0.027)	(0.007)	(0.025)	(0.308)
# observations	23,900	23,900	23,900	23,900	19,947	19,947	19,947	19,947
# est.	2,677	2,677	2,677	2,677	2,254	2,254	2,254	2,254
R-squared	0.921	0.923	0.883	0.899	0.921	0.924	0.885	0.906
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by county in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile and year in the top panel and by size quartile, county and year in the middle and lower panels. Observations are not weighted. All variables are winsorized at the first and 99th percentiles.

Alternative approaches to statistical inference. Tables D.8 and D.9 show that the effects in Table VII and Figure VIII are robust to clustering standard errors by firm and county. The tables do not include results for headquarters and establishments of firms with one establishment, because results are the same whether we cluster standard errors at the county or firm × county level, as units are nested within counties.

Tables D.10 and D.11 document that the main results of Table VII remain significant even if we adjust the hypotheses tests for possible multiple testing bias. We employ two different approaches to adjust for testing multiple outcomes and for testing multiple outcomes for multiple groups. The approach by Romano and Wolf (2005) allows us to take into account that we test multiple outcome variables for each group. The advantage of the approach is that it accounts for the correlation of variables, e.g., for the correlation between the number of managerial layers and the managerial share. Its disadvantage is that it is difficult to account for the fact that we test outcomes for multiple groups. We therefore also employ the approach by Holm (1979). The advantage of this approach is its simplicity: it is easy to account for multiple outcomes per group and for multiple groups. Its disadvantage is that it treats outcomes as independent, so it is too conservative.

Table D.10 displays the results of the Romano-Wolf approach. We also report the significance levels according to Holm (1979) applied by group in order to assess how much "too conservative" the latter approach is. Table D.11 displays the results of the Holm approach applied across outcomes and groups.

Table D.8: Table VII with standard errors clustered by firm × county, 2000–2010 panel

	All firms				Firms with ≥ 2 establishments			
Dep. variable	# em. (1)	wages (2)	# lay. (3)	mg.sh. (4)	# em. (5)	wages (6)	# lay. (7)	mg.sh. (8)
Directly affected establishment								
Est. treated	0.084*	0.002	0.009	0.059	0.083+	0.004	0.016	0.049
	(0.040)	(0.005)	(0.019)	(0.280)	(0.046)	(0.006)	(0.021)	(0.301)
# observations	47,732	47,732	47,732	47,732	40,143	40,143	40,143	40,143
# est.	5,609	5,609	5,609	5,609	4,791	4,791	4,791	4,791
R-squared	0.901	0.931	0.875	0.890	0.899	0.932	0.878	0.902
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.017	0.019**	0.020	0.521+
					(0.027)	(0.007)	(0.018)	(0.290)
# observations					45,508	45,508	45,508	45,508
# est.					5,508	5,508	5,508	5,508
R-squared					0.917	0.926	0.890	0.887
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by firm and county in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top panel and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Table D.9: Table D.4 with standard errors clustered by firm $\times$ county, 2000–2010 panel

	Directly affected establishments							
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alternative control groups								
	Sample w/o non-directly affected est.							
Est. treated	0.043^{+}	0.011^{*}	−0.003	0.281				
	(0.026)	(0.005)	(0.016)	(0.336)				
# observations	29,661	29,661	29,661	29,661				
# est.	3,336	3,336	3,336	3,336				
R-squared	0.921	0.921	0.873	0.863				
Est. FE	Y	Y	Y	Y				
County-year FE	Y	Y	Y	Y				
Entry before 2000								
	All firms				Firms with $\geq$ 2 establishments			
Est. treated	0.073^{*}	−0.001	−0.004	0.161	0.068^{*}	0.000	0.006	0.139
	(0.030)	(0.005)	(0.018)	(0.319)	(0.033)	(0.006)	(0.021)	(0.312)
# observations	41,255	41,255	41,255	41,255	34,807	34,807	34,807	34,807
# est.	4,666	4,666	4,666	4,666	3,986	3,986	3,986	3,986
R-squared	0.911	0.933	0.882	0.899	0.911	0.933	0.887	0.912
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Entry before 1995								
	All firms				Firms with $\geq$ 2 establishments			
Est. treated	0.085^{*}	−0.001	−0.016	−0.188	0.082^{+}	−0.002	−0.012	−0.204
	(0.042)	(0.007)	(0.023)	(0.389)	(0.049)	(0.008)	(0.028)	(0.399)
# observations	23,900	23,900	23,900	23,900	19,947	19,947	19,947	19,947
# est.	2,677	2,677	2,677	2,677	2,254	2,254	2,254	2,254
R-squared	0.918	0.926	0.885	0.904	0.918	0.926	0.890	0.917
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by firm and county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile and year in the top panel and by size quartile, county and year in the middle and bottom panels. All variables are winsorized at the first and 99th percentiles.

Table D.10: Adjusting hypotheses tests for multiple testing by group, 2000–2010 panel

	All firms				Firms with ≥ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.084	0.002	0.009	0.059	0.083	0.004	0.016	0.049
SE	0.019	0.004	0.016	0.253	0.021	0.005	0.018	0.273
P-value	0.000	0.637	0.591	0.817	0.000	0.476	0.387	0.859
Baseline α	***				***			
Romano-Wolf α	**				**			
Holm α	***				***			
Headquarters								
Firm treated	−0.013	0.004	0.019	0.103	0.023	0.018	0.065	0.994
SE	0.019	0.005	0.022	0.338	0.033	0.008	0.022	0.504
P-value	0.491	0.491	0.374	0.762	0.489	0.034	0.006	0.054
Baseline α						*	**	+
Romano-Wolf α						+	*	+
Holm α							*	
Non-directly affected establishment								
Firm treated					−0.017	0.019	0.020	0.521
SE					0.020	0.004	0.017	0.209
P-value					0.412	0.000	0.228	0.014
Baseline α						***		*
Romano-Wolf α						**		*
Holm α						***		*

Standard errors (SE) clustered by (headquarter) county and corresponding P-value. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Baseline α:* baseline significance level (see Table VII), *Romano-Wolf α:* significance level adjusted for multiple testing following Romano and Wolf (2005) computed using the `rwolf` command provided by Clarke et al. (2020), *Holm α:* significance level adjusted for multiple testing following Holm(1979) separately for establishments, headquarters and non-directly affected establishments. We bootstrap manually to be able to apply the `rwolf` command to our weighted baseline estimates (at the level of the (HQ) county for directly affected establishments and HQ, and at the level of the county x HQ county for non-directly affected establishments). *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Note: The P-value for the wages of headquarters is only marginally higher than the threshold required for significance at the 10% level (0.033874 > 0.033333)

Table D.11: Adjusting hypotheses tests for multiple testing across groups, 2000–2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.084	0.002	0.009	0.059	0.083	0.004	0.016	0.049
SE	0.019	0.004	0.016	0.253	0.021	0.005	0.018	0.273
P-value	0.000	0.637	0.591	0.817	0.000	0.476	0.387	0.859
Baseline α	***				***			
Holm α	***				**			
Headquarters								
Firm treated	−0.013	0.004	0.019	0.103	0.023	0.018	0.065	0.994
SE	0.019	0.005	0.022	0.338	0.033	0.008	0.022	0.504
P-value	0.491	0.491	0.374	0.762	0.489	0.034	0.006	0.054
Baseline α						*	**	+
Holm α							+	
Non-directly affected establishment								
Firm treated					−0.017	0.019	0.020	0.521
SE					0.020	0.004	0.017	0.209
P-value					0.412	0.000	0.228	0.014
Baseline α						***		*
Holm α						***		

Standard errors (SE) clustered by (HQ) county and corresponding P-value. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Baseline α:* baseline significance level (see Table VII), *Holm α:* significance level adjusted for multiple testing following Holm (1979) accounting for all hypotheses tested per sample. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Note: The P-value for the number of layers of headquarters is only slightly higher than the threshold required for significance at the 5% level (0.005702 > 0.005) and the P-value for the managerial share of non-directly affected establishments is only slightly higher than the threshold required for significance at the 10% level (0.014048 > 0.011).

Non-directly affected establishments: strata. Table D.12 documents that results for non-directly affected establishments are similar to the baseline results in Table VII if we match establishments by size quartile, county and year, or by size quartile, headquarter county and year.

Table D.12: Robustness to alternative strata, 2000–2010 panel

	Non-directly affected establishments							
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Strata: *county x size x year*				Strata: *HQ county x size x year*			
Firm treated	−0.007	0.021***	0.016	0.440+	−0.051+	0.008	0.014	0.671*
	(0.026)	(0.005)	(0.021)	(0.260)	(0.030)	(0.005)	(0.020)	(0.329)
# observations	34,007	34,007	34,007	34,007	36,835	36,835	36,835	36,835
# est.	4,382	4,382	4,382	4,382	4,690	4,690	4,690	4,690
R-squared	0.914	0.931	0.889	0.899	0.914	0.922	0.888	0.905
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. All variables are winsorized at the first and 99th percentiles.

D.5.2. Alternative variable definition

Alternative outcome variables. We use the total number of employees as size measure, average wages of all employees and the managerial span of control as alternative to non-managerial wages, and define the managerial share based on layers. To provide more direct evidence on employees' knowledge, we report results for the education and experience of the non-managerial employees and all employees. The results in columns (1)-(4) confirm our main regression results. The insignificant coefficients for the span of control likely reflect that our measure is coarse, because our data only contain the number of employees, but no information on the number of hours worked. The results for education and experience in columns (5)-(8) suggest that the increase of wages is indeed driven by higher knowledge of the employees, consistent with the mechanism in our model.

Table D.13: Lower travel times affect all units of ME firms, alternative outcomes, 2000–2010 panel

Dep. variable	# em. (1)	wages (2)	span (3)	mg.sh. (4)	ed. 0 (5)	ed. all (6)	ex. 0 (7)	ex. all (8)
Directly affected establishment								
Est. treated	0.084***	0.003	0.030	0.207	−0.006	0.001	−0.005	−0.015
	(0.021)	(0.005)	(0.029)	(0.338)	(0.012)	(0.011)	(0.019)	(0.016)
# observations	40,143	40,143	40,143	40,143	40,143	40,143	40,143	40,143
# est.	4,791	4,791	4,791	4,791	4,791	4,791	4,791	4,791
R-squared	0.905	0.945	0.833	0.914	0.839	0.886	0.753	0.765
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	0.043	0.019*	0.022	0.929	0.013	0.033*	−0.011	−0.012
	(0.027)	(0.008)	(0.054)	(0.616)	(0.018)	(0.013)	(0.033)	(0.031)
# observations	6,261	6,261	6,261	6,261	6,261	6,261	6,261	6,261
# HQ	683	683	683	683	683	683	683	683
R-squared	0.957	0.959	0.807	0.941	0.899	0.943	0.810	0.825
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated	−0.012	0.023***	0.030	0.344	−0.013	−0.008	0.027^{+}	0.027*
	(0.021)	(0.004)	(0.026)	(0.267)	(0.010)	(0.007)	(0.015)	(0.013)
# observations	45,508	45,508	45,508	45,508	45,508	45,508	45,508	45,508
# est.	5,508	5,508	5,508	5,508	5,508	5,508	5,508	5,508
R-squared	0.924	0.939	0.847	0.916	0.831	0.880	0.754	0.765
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log total number of employees; *wages:* average log wages of all employees; *span:* managerial span, defined as ratio of the number of employees over number of employees in the top layer (as in Friedrich 2020); *mg.sh.:* share of managerial occupations in wage sum in percent, where managerial occupations are determined by layer, *ed. 0, ex. 0:* education, average log experience (= # of days in establishment/HQ) of non-managerial employees, *ed. all, ex. all:* education, average log experience of all employees. Treated and control units are matched by size quartile, (headquarter) county and year, except for the non-directly affected establishments that are matched by size quartile and year. All variables are winsorized at the first and 99th percentiles. Sample of firms with at least two establishments.

Alternative definition of treatment. Table D.14 documents that the results are similar if we use a stricter and a broader approach to determine which establishments are directly affected by lower travel times and which establishments are only indirectly affected.

Columns (1)-(4) report the results of regressions that use a stricter definition. We consider establishments to be directly affected by lower travel times only if they have a direct connection to the headquarters (i.e., no changes) and the travel times to the headquarters decrease by at least 30 minutes. Results are very similar to our baseline results.

Columns (5)-(8) report the results of regressions that use a broader definition. We also consider establishments to be directly affected by lower travel times if their travel times to the headquarters do not decrease by at least 30 minutes, but the travel times of the closest other establishment do. Results are again very similar. As this alternative definition does not affect which headquarters are treated, we do not report the results for the headquarters.

Table D.14: Stricter and broader treatment definition, 2000-2010 panel

	Stricter definition				Broader definition			
Dep. variable	# em.	wages	span	mg.sh.	# em.	wages	span	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.070^{**}	0.002	0.007	0.007	0.084^{***}	0.001	0.016	−0.104
	(0.026)	(0.005)	(0.022)	(0.350)	(0.020)	(0.005)	(0.017)	(0.262)
# observations	29,659	29,659	29,659	29,659	44,516	44,516	44,516	44,516
# est.	3,453	3,453	3,453	3,453	5,276	5,276	5,276	5,276
R-squared	0.902	0.931	0.878	0.905	0.899	0.930	0.878	0.899
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	0.001	0.016^{*}	0.022	0.862^{+}				
	(0.044)	(0.008)	(0.028)	(0.457)				
# observations	5,641	5,641	5,641	5,641				
# HQ	613	613	613	613				
R-squared	0.949	0.945	0.875	0.921				
HQ FE	Y	Y	Y	Y				
HQ c.-year FE	Y	Y	Y	Y				
Non-directly affected establishment								
Firm treated	−0.029	0.016^{***}	0.007	0.191	−0.023	0.020^{***}	0.021	0.560^{*}
	(0.020)	(0.004)	(0.015)	(0.196)	(0.020)	(0.004)	(0.018)	(0.230)
# observations	49,571	49,571	49,571	49,571	43,759	43,759	43,759	43,759
# est.	5,946	5,946	5,946	5,946	5,306	5,306	5,306	5,306
R-squared	0.917	0.925	0.890	0.887	0.917	0.926	0.890	0.889
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, (headquarter) county and year, except for the non-directly affected establishments that are matched by size quartile and year. All variables are winsorized at the first and 99th percentiles.

Alternative layer definition. Table D.15 shows that the results are robust if we count the managerial layers at the establishment level. As we explain in section II.B, we treat the lowest layer in each firm as non-managerial in our baseline specification. The lowest layer need not be equal across all units of the firm, however. The robustness check shows that we obtain similar results if we treat the lowest layer in each establishment as non-managerial.

Table D.15: Regression results, establishment-level layer definition, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.088***	0.004	0.013	−0.049	0.085***	0.005	0.007	−0.076
	(0.021)	(0.005)	(0.016)	(0.273)	(0.025)	(0.006)	(0.018)	(0.268)
# observations	50,993	50,993	50,993	50,993	42,667	42,667	42,667	42,667
# est.	6,057	6,057	6,057	6,057	5,163	5,163	5,163	5,163
R-squared	0.893	0.920	0.862	0.891	0.893	0.921	0.867	0.898
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	0.009	0.005	0.028	0.122	0.039	0.018*	0.067*	0.839
	(0.024)	(0.006)	(0.020)	(0.349)	(0.040)	(0.009)	(0.025)	(0.565)
# observations	13,505	13,505	13,505	13,505	6,304	6,304	6,304	6,304
# HQ	1,478	1,478	1,478	1,478	686	686	686	686
R-squared	0.946	0.949	0.874	0.904	0.946	0.941	0.871	0.913
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.011	0.019***	0.024	0.652*
					(0.019)	(0.004)	(0.015)	(0.292)
# observations					48,529	48,529	48,529	48,529
# est.					5,987	5,987	5,987	5,987
R-squared					0.904	0.917	0.874	0.884
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Bundesland × year fixed effects. Table D.16 replaces the (headquarter) county-year fixed effects with Bundesland-year fixed effects. One may be concerned that the regressions "overfit" the data, because they contain both establishment (or headquarter) and (headquarter) county-year fixed effects. The Bundesland-year fixed effects reduce the number of spatial fixed effects substantially from up to 1,200 to less than 180. The estimated effects are very similar.

Table D.16: Regression results, Bundesland-year fixed effects, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.084***	0.002	0.012	0.083	0.084***	0.004	0.020	0.105
	(0.019)	(0.004)	(0.016)	(0.246)	(0.020)	(0.005)	(0.018)	(0.270)
# observations	47,732	47,732	47,732	47,732	40,143	40,143	40,143	40,143
# est.	5,609	5,609	5,609	5,609	4,791	4,791	4,791	4,791
R-squared	0.898	0.929	0.871	0.887	0.895	0.929	0.874	0.899
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.015	0.003	0.024	0.149	0.019	0.017*	0.066**	0.967+
	(0.019)	(0.005)	(0.021)	(0.324)	(0.034)	(0.008)	(0.023)	(0.509)
# observations	13,393	13,393	13,393	13,393	6,261	6,261	6,261	6,261
# HQ	1,469	1,469	1,469	1,469	683	683	683	683
R-squared	0.949	0.948	0.867	0.909	0.949	0.942	0.863	0.914
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.029	0.013***	0.006	0.560**
					(0.018)	(0.004)	(0.014)	(0.205)
# observations					45,658	45,658	45,658	45,658
# est.					5,522	5,522	5,522	5,522
R-squared					0.903	0.918	0.883	0.881
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

D.5.3. Alternative sample restrictions

Omit route 2. As explained in Appendix D.2, routes 1 and 3 are exclusively used by passenger trains, whereas route 2 is also used by freight trains. To ensure that our results are not driven by changes in transport costs, we replicate our results, considering as treated only establishments with travel time reductions due to the opening of routes 1 and 3. Table D.17 shows that our results are robust, although the effect on the number of managerial layers is now only marginally significant (P-value 11.7%). The results are less precisely estimated, reflecting that it adds noise to the data to consider units untreated that are affected by lower travel times after the opening of route 2.

Table D.17: Regression results without route 2, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.094***	0.004	0.012	0.005	0.090***	0.007	0.020	0.015
	(0.023)	(0.005)	(0.018)	(0.264)	(0.023)	(0.006)	(0.020)	(0.286)
# observations	44,808	44,808	44,808	44,808	37,534	37,534	37,534	37,534
# est.	5,281	5,281	5,281	5,281	4,499	4,499	4,499	4,499
R-squared	0.900	0.930	0.875	0.893	0.897	0.930	0.878	0.904
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.017	0.003	0.007	0.137	−0.002	0.019*	0.040	1.026+
	(0.020)	(0.006)	(0.024)	(0.377)	(0.034)	(0.009)	(0.025)	(0.551)
# observations	13,153	13,153	13,153	13,153	6,235	6,235	6,235	6,235
# HQ	1,440	1,440	1,440	1,440	680	680	680	680
R-squared	0.952	0.953	0.874	0.918	0.951	0.945	0.870	0.928
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.013	0.022***	0.021	0.496*
					(0.022)	(0.004)	(0.016)	(0.220)
# observations					46,596	46,596	46,596	46,596
# est.					5,613	5,613	5,613	5,613
R-squared					0.917	0.926	0.890	0.887
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Drop largest units/ stations located on the route. One could be concerned that firms invested in lobbying activities to influence the course of the routes, which may bias our coefficients. We take two approaches to address this concern. First, we follow Charnoz et al. (2018) and exclude the largest establishment or headquarters in each county from the sample. Second, we exclude establishments or headquarters in counties served by train stations located on the route.[18] During the planning process, usually several possible courses to connect the two endpoints of the routes were considered. Establishments and headquarters in counties served by stations on the route therefore should have the highest lobbying incentives. Tables D.18 and D.19 show that our results are robust to both modifications. This suggests that our results are unlikely to be driven by lobbying.

Table D.18: Regression results, excluding largest units per county, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.085***	0.002	0.006	0.029	0.073**	0.005	0.012	0.032
	(0.020)	(0.004)	(0.016)	(0.261)	(0.022)	(0.005)	(0.018)	(0.267)
# observations	46,534	46,534	46,534	46,534	38,663	38,663	38,663	38,663
# est.	5,508	5,508	5,508	5,508	4,633	4,633	4,633	4,633
R-squared	0.896	0.930	0.872	0.891	0.894	0.931	0.874	0.902
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.013	0.002	0.028	0.267	0.032	0.020*	0.072**	1.072*
	(0.021)	(0.005)	(0.021)	(0.308)	(0.035)	(0.009)	(0.025)	(0.512)
# observations	12,529	12,529	12,529	12,529	5,768	5,768	5,768	5,768
# HQ	1,395	1,395	1,395	1,395	641	641	641	641
R-squared	0.946	0.947	0.870	0.913	0.950	0.944	0.872	0.919
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year. All variables are winsorized at the first and 99th percentiles.

[18] Specifically, we exclude the stations Limburg Süd, Montabaur and Siegburg/Bonn located on route 1, and Ludwigslust and Wittenberge located on route 2. We exclude Ingolstadt even though it is strictly speaking an endpoint of route 3, because route 3 was built to better connect Munich and Nuremberg, and Ingolstadt competed with Augsburg for the route. Finally, for consistency, we exclude Lutherstadt Wittenberg located on the route between Leipzig and Berlin, even though we cannot use this route in our estimation (see footnote 11).

Table D.19: Regression results, excluding stations on route, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.082***	0.002	0.008	0.080	0.084***	0.003	0.017	0.059
	(0.019)	(0.004)	(0.017)	(0.270)	(0.021)	(0.005)	(0.019)	(0.289)
# observations	46,143	46,143	46,143	46,143	38,772	38,772	38,772	38,772
# est.	5,381	5,381	5,381	5,381	4,585	4,585	4,585	4,585
R-squared	0.899	0.928	0.871	0.892	0.898	0.928	0.875	0.904
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.018	0.005	0.021	0.203	0.028	0.014	0.074***	1.146*
	(0.020)	(0.005)	(0.022)	(0.336)	(0.034)	(0.009)	(0.021)	(0.551)
# observations	13,015	13,015	13,015	13,015	6,201	6,201	6,201	6,201
# HQ	1,418	1,418	1,418	1,418	679	679	679	679
R-squared	0.952	0.951	0.874	0.913	0.952	0.944	0.872	0.913
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year. All variables are winsorized at the first and 99th percentiles.

Drop moving establishments. Table D.20 replicates the regressions after dropping establishments or headquarters that move between counties. The results for the establishments are virtually unchanged; the results for the headquarters even become stronger.

Table D.20: Regression results, without moving establishments/ headquarters, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.082***	0.001	0.006	0.170	0.081***	0.003	0.012	0.057
	(0.019)	(0.005)	(0.017)	(0.263)	(0.019)	(0.006)	(0.019)	(0.282)
# observations	44,357	44,357	44,357	44,357	36,996	36,996	36,996	36,996
# est.	5,230	5,230	5,230	5,230	4,409	4,409	4,409	4,409
R-squared	0.903	0.932	0.879	0.895	0.901	0.932	0.881	0.906
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.021	0.004	0.019	0.080	0.013	0.016*	0.075**	1.030+
	(0.025)	(0.005)	(0.021)	(0.289)	(0.042)	(0.007)	(0.027)	(0.552)
# observations	12,587	12,587	12,587	12,587	5,847	5,847	5,847	5,847
# HQ	1,382	1,382	1,382	1,382	652	652	652	652
R-squared	0.954	0.955	0.885	0.921	0.953	0.951	0.889	0.918
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.018	0.020***	0.019	0.418*
					(0.021)	(0.004)	(0.017)	(0.207)
# observations					43,007	43,007	43,007	43,007
# est.					5,236	5,236	5,236	5,236
R-squared					0.918	0.925	0.893	0.892
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

Only always connected units. Table D.21 shows that the results are robust to restricting the sample to establishments and headquarters that are connected to the long-distance network in all years. This ensures that the high-speed routes, not (dis)connecting stations to the network, drive changes in travel time.

Table D.21: Regression results, 2000-2010 panel, only always connected stations

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.081***	0.002	0.003	−0.150	0.085**	0.000	0.011	−0.144
	(0.023)	(0.005)	(0.022)	(0.307)	(0.028)	(0.006)	(0.025)	(0.332)
# observations	33,247	33,247	33,247	33,247	28,455	28,455	28,455	28,455
# est.	3,849	3,849	3,849	3,849	3,353	3,353	3,353	3,353
R-squared	0.902	0.922	0.869	0.902	0.902	0.921	0.871	0.909
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.018	0.004	0.007	0.102	0.003	0.015	0.047*	1.255*
	(0.020)	(0.006)	(0.026)	(0.358)	(0.029)	(0.010)	(0.022)	(0.517)
# observations	10,239	10,239	10,239	10,239	4,970	4,970	4,970	4,970
# HQ	1,104	1,104	1,104	1,104	543	543	543	543
R-squared	0.948	0.945	0.869	0.916	0.954	0.935	0.871	0.921
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.021	0.017***	0.016	0.548*
					(0.025)	(0.004)	(0.021)	(0.262)
# observations					30,724	30,724	30,724	30,724
# est.					3,654	3,654	3,654	3,654
R-squared					0.921	0.923	0.889	0.896
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year in the top and middle panels and by size quartile and year in the bottom panel. All variables are winsorized at the first and 99th percentiles.

D.6. Effect heterogeneity depending on the predictability of the production process λ

Lower travel times make it easier for headquarter managers to travel to the establishments. The model spells out a specific channel how this reduction of within-firm geographic frictions affects firm organization, namely through lower helping costs between the headquarters and the establishment. At the same time, lower travel times may reduce other managerial frictions, such as monitoring costs (Giroud 2013).

To support the helping cost channel, we document that the estimated effects are heterogeneous across sectors. We use a sector-level measure for λ, the predictability of the production process. A higher value of λ means that problems in the tail of the problem probability distribution occur with lower probability. The higher λ, the higher is the reduction in the number of problems sent to the CEO that is caused by an increase in local knowledge. Changes in helping costs should therefore have more pronounced effects on the endogenous choices in sectors with a less predictable production process, i.e., lower value of λ. Figure D.1 contains the results of simulations corroborating this heterogeneity.

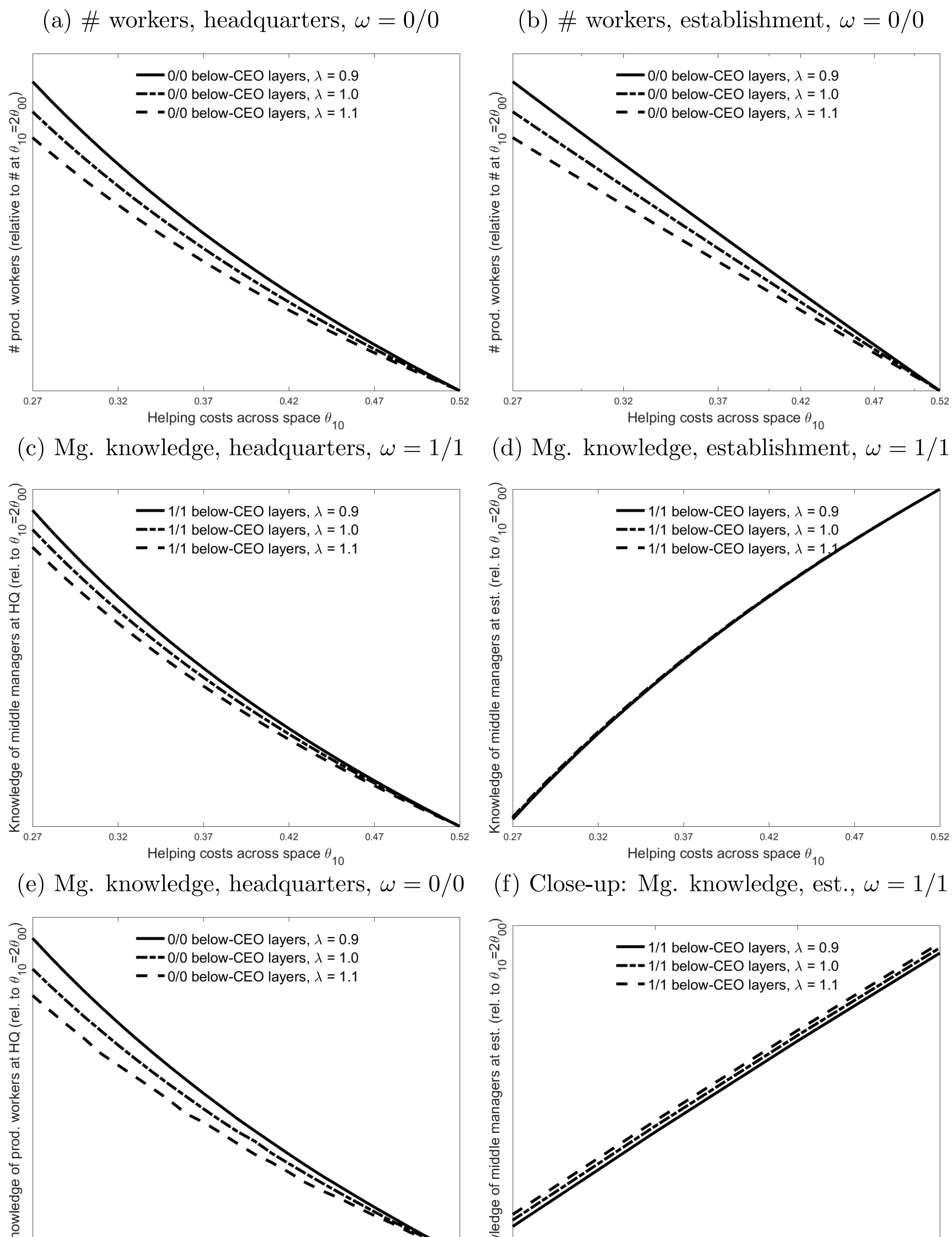
We use the measure of the predictability of the production process from Gumpert (2018). The measure is based on a survey question from the "BiBB/BAuA Employment Survey 2006" (Hall and Tiemann 2006). The survey provides data on the employment conditions of a representative sample of 20,000 workers in Germany. The measure exploits the question of how often respondents have "to react to and solve unforeseeable problems" in their current job. It is constructed by restricting the sample to two-digit sectors with at least 25 respondents and regressing a dummy that is equal to one if participants answer "frequently," and zero if they answer "sometimes" or "never," on sector dummies. The estimated coefficients of the sector dummies are inversely related to the predictability of the production process.

We merge the measure to our data using the headquarter sector of a firm. As the measure of predictability is not available for all sectors in our sample, we end up with a lower number of units than in Table VII. We separately run the regressions for firms with above- and below-median predictability. We focus on the full sample of firms and, correspondingly, on the results for establishments and headquarters. If we restrict the sample to firms with at least two establishments, the sample size decreases excessively for meaningful headquarter regressions.

Table D.22 documents that firms in sectors with below-median predictability of the production process drive the results in Table VII. As a comparison of columns 1 to 4 and columns 5 to 8 shows, treated establishments in those sectors grow significantly faster than untreated ones. The effect size is about 50% higher than in the full sample. In contrast, the treated establishments in sectors with above-median predictability grow at the same rate as untreated ones. Interestingly, we find that headquarters in the below-median predictability sample increase their number of managerial layers, but headquarters in the above-median predictability sample increase wages. The positive effect on the number of layers is more pronounced than the corresponding effect in Table VII (both for the full sample and the sample of firms with at least two establishments). In contrast, the effect on wages tends is less pronounced.

While Table D.22 does not refute the possibility that lower travel times reduce monitoring costs within firms, it corroborates that the model mechanism of lower helping costs across space is at play in driving the effects documented in Table VII.

Figure D.1: Heterogeneity of the effect of θ_{10} depending on the predictability λ



(a) # workers, headquarters, $\omega = 0/0$

(b) # workers, establishment, $\omega = 0/0$

(c) Mg. knowledge, headquarters, $\omega = 1/1$

(d) Mg. knowledge, establishment, $\omega = 1/1$

(e) Mg. knowledge, headquarters, $\omega = 0/0$

(f) Close-up: Mg. knowledge, est., $\omega = 1/1$

The figure plots the number and knowledge of production workers and middle managers of a ME firm as a function of the helping costs across space $\theta_{10} \in [.27, .52]$ for different values of λ with $\omega \in \{0/0, 1/1\}$, $\tau = 1.1$, $w_1 = w_0$, and $\tilde{q}_1 = \tilde{q}_0 = 50$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg 2012), $w_0 = 1$. As the effect of λ relative to the effect of θ_{10} is small for the managerial knowledge at the establishment, Figure D.1f provides a close-up of the simulation for small values of θ_{10}.

Table D.22: Regression results, sample split by predictability, 2000–2010 panel

	Below-median predictability				Above-median predictability			
Dep. variable	# em.	wages	# lay.	mg.sh.	# em.	wages	# lay.	mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.133***	−0.001	0.015	0.219	0.045	0.010	0.009	0.004
	(0.031)	(0.006)	(0.034)	(0.376)	(0.027)	(0.007)	(0.019)	(0.301)
# observations	18,400	18,400	18,400	18,400	20,927	20,927	20,927	20,927
# est.	2,290	2,290	2,290	2,290	2,311	2,311	2,311	2,311
R-squared	0.892	0.922	0.848	0.942	0.916	0.933	0.897	0.788
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	0.009	0.000	0.099*	0.902	−0.027	0.014+	−0.037	−0.318
	(0.064)	(0.007)	(0.041)	(0.681)	(0.030)	(0.007)	(0.026)	(0.292)
# observations	5,545	5,545	5,545	5,545	5,529	5,529	5,529	5,529
# HQ	611	611	611	611	615	615	615	615
R-squared	0.944	0.948	0.881	0.926	0.961	0.951	0.874	0.874
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Standard errors clustered by (HQ) county in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table VII. Treated and control units are matched by size quartile, county and year. All variables are winsorized at the first and 99th percentiles.

Appendix References

Antoni, Manfred, Andreas Ganzer, and Philipp vom Berge (2016). Sample of integrated labour market biographies (SIAB) 1975-2014. FDZ-Datenreport 04/2016, Institute for Employment Research.

Antoni, Manfred, Katharina Koller, Marie-Christine Laible, and Florian Zimmermann (2018). Orbis-ADIAB: From record linkage key to research dataset. Combining commercial company data with administrative employer-employee data. FDZ-Methodenreport 04/2018, Institute for Employment Research.

Antoni, Manfred, and Rainer Schnell (2019). The past, present and future of the German Record Linkage Center (GRLC). *Jahrbücher für Nationalökonomie und Statistik* 239(2), 1–13.

Belter, Bringfried (2004). Qualitätssicherung im Bereich Ingenieurleistungen im Bahnbau. *EIK Eisenbahn Ingenieur Kompendium*, 249–256.

Blossfeld, Hans-Peter (1983). Höherqualifizierung und Verdrängung—Konsequenzen der Bildungsexpansion in den siebziger Jahren. In Haller, Max and Müller, Walter (Eds.), *Beschäftigungssystem im gesellschaftlichen Wandel*, pp. 184-241. Frankfurt/New-York: Campus-Verlag.

Blossfeld, Hans-Peter (1987). Labor-market entry and the sexual segregation of careers in the Federal Republic of Germany. *American Journal of Sociology* 93 (1), 89-118.

Brux, Gunther (2003). Tunnelbau für die Neubaustrecke Köln-Rhein/Main: Spritzbetonbauweise und Weiterentwicklungen. *EIK - Eisenbahn Ingenieur Kompendium*, 289-308.

Brux, Gunther (2006). NBS Nürnberg - Ingolstadt in Betrieb. *Eisenbahningenieur* 57 (6), 16-21.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg (2015). The Anatomy of French Production Hierarchies. *The Journal of Political Economy* 123 (4), 809-852.

Caliendo, Lorenzo and Esteban Rossi-Hansberg (2012). The impact of trade on organization and productivity. *The Quarterly Journal of Economics* 127 (3), 1393-1467.

Charnoz, Pauline, Claire Lelarge, and Corentin Trevien (2018). Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail. *The Economic Journal* 128 (610), 949-994.

Christen, Peter (2012). *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection.* Berlin: Springer.

Deutscher Bundestag (2007). Flughafenanbindungen von Fernverkehrshaltepunkten der Bahn. BT-Drucksache 16/7104.

Deutscher Bundestag (2010). Schienengüterverkehr auf den Neubaustrecken Nürnberg–Ingolstadt und Wendlingen–Ulm. BT-Drucksache 17/3311.

Eberle, Johanna, Peter Jacobebbinghaus, Johannes Ludsteck, and Julia Witter (2011). Generation of time-consistent industry codes in the face of classification changes: Simple heuristic based on the Establishment History Panel (BHP). FDZ-Methodenreport 05/2011, Institute for Employment Research.

Eurailpress.de (2007). Schnelle Bahnverbindung: Lufthansa streicht Flüge Köln Frankfurt. http://www.eurailpress.de/news/alle-nachrichten/single-view/news/schnellebahnverbindung-lufthansa-streicht-fluege-koeln-frankfurt-1.html, accessed on 24 January 2018.

Feldwisch, Wolfgang, Olaf Drescher, and Christine Haag (2004). Tempo 230 zwischen Hamburg und Berlin. *Eisenbahntechnische Rundschau*, 821-831.

Feldwisch, Wolfgang and Holger Schülke (2006). Die Inbetriebnahme der Großprojekte der Bahn zur Fuballweltmeisterschaft 2006. *Eisenbahntechnische Rundschau*, 289-300.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter (2005). Imputation rules to improve the education variable in the IAB employment subsample. *Schmollers Jahrbuch* 126 (3), 405436.

Friedrich, Benjamin U. (2020). Trade Shocks, Firm Hierarchies, and Wage Inequality. *Review of Economics and Statistics*, forthcoming.

Giroud, Xavier (2013). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics* 128 (2), 861-915.

Gumpert, Anna (2018). The organization of knowledge in multinational firms. *Journal of the European Economic Association* 16 (6), 1929-1976.

Hall, Anja and Michael Tiemann (2006). BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany 2006, suf 1.0. Technical report, GESIS Cologne, Germany (data access); Federal Institute of Vocational Education and Training, Bonn doi:10.4232/1.4820.

Hethey, Tanja and Johannes Schmieder (2010). Using Worker Flows in the Analysis of Establishment Turnover Evidence from German Administrative Data. FDZ-Methodenreport 06/2010.

Holm, Sture (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Iacus, Stefano M., Gary King, and Giuseppe Porro (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis* 20 (1), 1-24.

Jacobebbinghaus, Peter and Stefan Seth (2007). The German integrated employment biographies sample IEBS. *Schmollers Jahrbuch* 127 (2), 335-342.

Jänsch, Eberhard (2006). Hochgeschwindigkeitsverkehr in Deutschland - 15 Jahre Erfolg. *Eisenbahntechnische Rundschau*, 704-710.

Jänsch, Eberhard (2010). Ab 1991: Schnelle Züge für den Fernverkehr. *Eisenbahntechnische Rundschau*, 36-41.

Jänsch, Eberhard (2015). Mischverkehr auf Neubaustrecken. *Eisenbahntechnische Rundschau*, 28-32.

Kaniut, Peter and Egbert Form (2002). Die Rolle des EBA bei Großprojekten der DB AG am Beispiel der NBS Köln-Rhein/Main. *Eisenbahntechnische Rundschau*, 554564.

Romano, Joseph P., and Michael Wolf (2005). Stepwise multiple testing as formalized data snooping. *Econometrica* 73, 1237–1282.

Schild, Christopher-Johannes (2016). Linking "Orbis Company" Data with Establishment Data from the German Federal Employment Agency. German RLC Working Paper No. wp-grlc-2016-02, http://grlc.german-microsimulation.de/wp-content/uploads/2017/07/wp-grlc-2016-02.pdf.

Schmucker, Alexandra, Stefan Seth, Johannes Ludsteck, Johanna Eberle, and Andreas Ganzer

(2016). The Establishment History Panel 1975-2014. FDZ-Datenreport 03/2016.

Tåg, Joacim (2013). Production hierarchies in Sweden. *Economics Letters* 121 (2), 210-213.

Wegener, Peter (2003). Bewältigung der Karstproblematik bei der NBS NürnbergIngolstadt. *Eisenbahntechnische Rundschau*, 182-191.

Weigelt, Horst (2003). Ursprung, Funktion und Nutzen der Neubau-/Ausbaustrecke Nürnberg–München. *Eisenbahntechnische Rundschau*, 192-201.