Firm organization with multiple establishments

Anna Gumpert* Henrike Steimer** Manfred Antoni*⁎

November 2018

Abstract

How do geographic frictions affect firm organization? We show theoretically and empirically that geographic frictions increase the use of middle managers in multi-establishment firms. In our model, we assume that the time of the CEO of a firm is a resource of limited supply that is shared among the headquarters and the establishments. Geographic frictions increase the costs of accessing the CEO. Hiring middle managers at an establishment substitutes for CEO time that is reallocated over all establishments. In consequence, 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 train 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.

*LMU Munich, Seminar for Comparative Economics, Akademiestr. 1, 80799 Munich, Germany, CESifo, and CEPR, e-mail: anna.gumpert@econ.lmu.de.

**Stanford Institute for Economic Policy Research, John and Cynthia Fry Gunn Building, 366 Galvez Street, Stanford, CA 94305, USA, e-mail: hsteimer@stanford.edu.

*⁎Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB), Regensburger Strasse 104, 90478 Nuremberg, Germany.

A prior version of this paper circulated under the title "Communication and the organization of firms across space".

1 Introduction

Large firms often organize their employees in multiple establishments at different locations. Within multi-establishment firms, geographic frictions such as long travel times to the headquarters adversely affect establishment performance (e.g., Giroud, 2013). Anecdotal evidence suggests that adjusting the managerial organization may help firms mitigate the negative impact of geographic frictions. For example, employing middle managers at regional offices instead of the headquarters was a key ingredient for the success of Singer Sewing Machine in the US (Chandler, 2002, p. 403-5). Philips employed dedicated country managers and regional executives as part of a larger strategy to revitalize their operations in the 1990s (Nueno and Ghemawat, 2002). And when the Canadian firm Blinds To Go set up a manufacturing plant in New Jersey in 1998, moving an experienced manager on site proved vital to improve the new plant's production efficiency (Menor and Mark, 2001). Yet, to date, there is no theoretical work and only very little empirical evidence on the impact of geographic frictions on the managerial organization of firms.

This paper studies the managerial organization of firms with multiple establishments. We show empirically and theoretically that geographic frictions increase the use of middle managers in multi-establishment firms. We use a new data set from administrative sources in Germany to document that multi-establishment firms with more distant establishments employ more employees in managerial occupations. In our model, geographic frictions increase the optimal number of managerial layers of multi-establishment firms and affect the optimal organization of the distant establishment. Importantly, the organizational adjustments at the establishment have repercussions for the managerial organization of the headquarters. Hence, the model predicts that geographic frictions between the headquarters and one establishment affect the optimal managerial organization of the headquarters and possible other establishments of a multi-establishment firm. We use our data to show that this prediction is reflected in the organizational response of multi-establishment firms to a reduction of travel times after the opening of high-speed train routes.

A key implication of our study is that the managerial organization of firms with multiple establishments is interdependent across establishments. This implies that local economic conditions not only affect the local establishment, but also the headquarters and the other establishments of a multi-establishment firm. Local conditions thus propagate across space through firm organization.

We motivate our study by documenting three stylized facts about multi-establishment firm organization. Our data set is ideally suited to study multi-establishment firms because it combines detailed information about the employees and the location of the establishments of a firm. The facts suggest that geographic frictions affect the location and organization of multi-establishment firms.

First, the probability that a firm operates an establishment at a locality decreases with the distance of the locality from its headquarters. Distance from the headquarters also correlates negatively with the size of the establishments.

Second, the number of managerial layers of multi-establishment firms correlates positively with the distance of its establishments from the headquarters. The correlation is neither driven by larger firms investing at farther locations nor by other firm characteristics. Quantitatively, doubling the distance is associated with the same increase of the number of layers as increasing sales by 14 percent. Distance correlates positively with the number of managerial layers both at the establishments and the headquarters.

Third, multi-establishment firms reorganize gradually as they grow. They typically add or drop managerial layers either at the headquarters or the establishments. Only rarely, they change the number of layers at both headquarters and establishments at the same time. This pattern is similar for firms with few and many establishments, and for firms with close and distant establishments.

We propose a model to understand how geographic frictions affect the optimal managerial organization of firms. We assume that a firm consists of a headquarters and possibly an additional establishment. The production workers in the headquarters and the establishment share a common CEO. The CEO is located at the headquarters and helps the workers solve problems that arise in the production process. Production is a problem solving process (as in Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). Workers input labor and generate problems that have to be solved using their knowledge or the knowledge of the CEO to generate output. The firm may choose to hire a layer of local middle managers. The middle managers solve part of the problems that would otherwise have to be solved by the CEO, but entail a quasi-fixed cost for the firm.

Helping workers costs CEO time. The driving forces behind the theoretical results 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 distant establishment.

As they strain CEO time, geographic frictions decrease the probability that a firm operates an establishment. For the same reason, establishments are typically smaller than the headquarters. This result is consistent with the lower investment probability at distant locations and the lower size of distant establishments documented in Fact 1.

Through limited CEO time, geographic frictions affect the organization of both the establishment and the headquarters. The firm adjusts the establishment organization in response to more severe geographic frictions so that fewer problems have to be solved by the CEO. In particular, geographic frictions make it more useful to hire middle managers. As the CEO is shared among the headquarters and the establishment, the firm also adjusts the organization at the headquarters. The model thus explains Fact 2: the number of layers increases with geographic frictions in the data, and the 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 increases efficiency also at the headquarters (and vice versa). The reason is that middle managers release CEO time, so middle managers at the establishment increase the amount of CEO time available for the headquarters. As a result, they decrease the need for hiring middle managers at the headquarters. This result explains Fact 3 that multi-establishment firms do not add layers at the headquarters and the establishments at the same time as they grow. Both the gradual reorganization and the impact of geographic frictions reflect that multi-establishment firm organization is interdependent across establishments.

In the final part of our paper, we exploit the opening of high-speed train routes in Germany to study the response of firm organization to exogenous variation in geographic frictions. The train routes reduce the travel time between the establishments and the headquarters. The new connections provide the fastest mode of travel between locations: they are faster than cars or planes (if one accounts for waiting times at the airport). We focus on the model prediction that geographic frictions between the headquarters and one establishment have repercussions for the organization of the headquarters and possible other establishments of the firm. Importantly, geographic frictions affect the investment probability and firm size in the model. They therefore have an indirect effect through size on the managerial organization in addition to their direct effect. Changes of output lead to changes in the number of workers and thus to changes in the number of layers. Only the total—direct and indirect—effect of lower travel times is identified.

We find that establishments benefiting from faster travel times grow faster than other establishments. Their number of managerial layers does not change significantly. 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 outweighing each other. Importantly, we find that faster travel times increase the number of managerial layers at the headquarters even though headquarter size does not change. This finding supports the interdependence of establishment organization predicted by the model. The interdependence goes beyond the headquarters if firms have multiple establishments one of which is affected by faster travel times: not directly affected establishments of firms with at least one establishment with faster travel times grow more slowly than other establishments, but their share of employees in managerial occupations tends to increase faster.

Our paper contributes to several strands of the literature. The key insight of the paper is that multi-establishment firm organization is interdependent across establishments. This insight is particularly relevant for a recent literature that documents that multi-establishment firms propagate local shocks through their internal networks (Giroud and Mueller, 2017; Seetharam, 2018). The literature discusses managerial and financial constraints as possible drivers of the empirical findings. However, although CEOs are considered an important determinant of firm performance (see Bertrand, 2009, for a survey), managerial constraints have received only very little systematic attention so far. Our contribution is to provide both a formal analysis and detailed empirical evidence on the role of managerial constraints for multi-establishment firm organization.

The interdependence of establishment organization is also relevant for the literature on multinational firms. In this 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 the public good assumption may apply to patents or trademarks, but not necessarily to managerial inputs.

Our paper contributes to the literature about the impact of distance to the headquarters and other geographic frictions on establishment performance in multi-establishment firms (e.g., Giroud, 2013; Kalnins and Lafontaine, 2013).[1] Our paper is particularly close to Charnoz et al. (2015). To our knowledge, they provide the only and purely empirical study of the impact of geographic frictions on firm organization. They focus on business groups and also exploit the opening of high-speed train routes. Our theoretical model shows why the impact of geographic frictions goes beyond a particular establishment, as also found by Charnoz et al. (2015), and cleanly disentangles the direct effects of geographic frictions on firm organization and the indirect effects through firm size. Our empirical analysis focuses on multi-establishment firms, i.e., the establishments are not legally independent units.

Our empirical strategy builds on the literature that uses the introduction of high-speed train routes to identify the impact of reductions of geographic frictions on firm outcomes (e.g. Bernard et al., 2017; Charnoz et al., 2015). To develop our model, we build on the literature of firms as knowledge hierarchies (for an overview, see Garicano and Rossi-Hansberg, 2015). A series of papers formalizes the idea that adding a layer of middle managers allows firms to increase efficiency as they grow, and assemble empirical evidence consistent with this hypothesis (e.g., Caliendo and Rossi-Hansberg, 2012; Caliendo et al., 2015a, b; Friedrich, 2016).[2] Similar

[1] Battiston et al. (2017) show that frictions to face-to-face communication decrease productivity in teams.

[2] Mariscal (2018) shows that the impact of new information technologies on firm organization explains the decline of the labor share in the U.S. using a knowledge hierarchy framework. Sforza

theoretical predictions result from a monitoring hierarchy framework (e.g., Chen, 2017; Chen and Suen, 2017).[3] The literature focuses on size as main determinant of organization. The possibility of multi-establishment production is largely neglected, even though multi-establishment firms are among the largest firms in developed economies and account for a substantial share of aggregate employment.[4]

Finally, our paper offers a novel perspective on the recent management literature. Bloom et al. (2016) document that half of the total variation in management practices between different U.S. establishments is due to variation between establishments within the same firm. Implementing managerial practices requires managerial time. The heterogeneity of management practices in multi-establishment firms may reflect asymmetries in the number of managerial layers and the amount of CEO time allocated to an establishment.

The paper is structured as follows. Section 2 describes the data. Section 3 presents the facts on multi-establishment firm organization. Section 4 develops the model. Section 5 presents the evidence from the opening of high speed train routes. The last section concludes.

2 Data

2.1 Data sources

We use a detailed linked firm-establishment-employee data set for Germany that is uniquely suited to study multi-establishment firms. The data contain information on the sales and the legal form of firms, and the location at the county level and the sector of their establishments.[5] We observe all employees per establishment subject to social security contributions on 30 June. The data include the occupation, level of education, employment history and wages of each employee. The data cover all sectors and the period 2000-2012. Each employee, establishment and firm has a unique identifier that allows following 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 (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) makes the data available for research. We use the employee history, the Establishment History Panel and the extension file entry

(2017) studies the organizational responses of firms to a credit supply shock.

[3]In the empirical literature on firm hierarchies, Rajan and Wulf (2006) document the flattening of corporate hierarchies over time, and Guadalupe and Wulf (2010) study the impact of competition on corporate hierarchies using detailed data on 300 large publicly traded U.S. firms.

[4]Gumpert (2018) contains 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.

[5]German counties are roughly comparable to counties in the USA.

and exit.[6] The Orbis database of Bureau van Dijk (BvD) contains balance sheet information of firms. We combine the social security records and the Orbis database using record linkage techniques. The algorithm exploits the regulation that the establishment names in the social security data have to contain the firm name. We identify the headquarters (HQ) establishment of a firm as the establishment with the same zip code or locality as the firm.[7] 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 use the 2012 cross section for cross-sectional analyses, because it contains the maximum number of establishments. The panel analyses use the period 2000-2010. We exclude the year 2011 because of changes in the occupational classification in that year (for details, see Appendix A.2). Consistent with the literature, we restrict our sample to full-time employees (e.g., Card et al., 2013; Dustmann et al., 2009). We focus on firms with at least 10 employees in all years. 99% of the firms dropped due to this requirement are small firms with only one establishment.

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

2.2 Measures for the 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., 2015b):

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 transfer the mapping in Caliendo et al. (2015b) based on the French classification of occupations to the German classification using official correspondence tables (Friedrich, 2016, uses an analogous procedure for Danish data). We treat the layer at the lowest level in each establishment as non-managerial. We count the number of layers above the lowest layer per firm. The lowest layer contains employees at level 0 in 98 percent of firms. Multi-establishment firms may separate management

[6] The establishment identifier in the Establishment History Panel may change when firm ownership changes. The extension file entry and exit allows following the establishments nonetheless.

[7] The social security data contain the address of an establishment. We are allowed to use the address to identify headquarters, but not to conduct our analyses for confidentiality.

Table 1: Descriptive statistics, SE vs. ME firms, 2012 cross section

Units of observation	N	% share ME firms						
Firms	109,357	9.0						
with non-missing sales	57,811	9.2						
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	1980	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

Descriptive statistics. *ME:* indicator for multi-establishment firm; *# employees:* number of full-time employees; *Sales (M €):* sales in million €.

and production, which is why we cross-check our findings treating the lowest level in the firm as non-managerial. Appendix A.3 provides details on our procedure and a list of occupations by level.

Alternatively, we use shares of managerial occupations in the wage sum. 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. The establishments report the occupations of the employees in the social security data. If a firm has several establishments, establishments may assign different occupations to similar employees. Cross-checking the results on the number of layers with the managerial share helps ensure that our results are robust to this possibility. On the other hand, we use the assignment of Blossfeld (1983, 1987, see Appendix A.3 for the list of managerial occupations). It builds on research from sociology and is part of Establishment History Panel. Managers are employees in occupations that have control or decision-making power over the use of production factors as well as high-level officials in organizations (Blossfeld, 1983, p. 208).[8]

Appendix A.4 illustrates the plausibility of the assignment of employees to layers. We show that employees at higher layers earn higher wages and have higher levels of education in the social security data. Further, we document that the tasks of employees systematically differ between layers in ways that plausibly reflect different roles of employees within firms using additional survey data.

2.3 Descriptive statistics

Table 1 provides descriptive statistics of the 2012 cross-section. Our sample comprises 109 thousand firms. We do not observe sales for all firms, but only the larger

[8]The German social security data do not contain the number of hours worked, so it is difficult to construct other measures commonly used in the literature such as the span of managerial control.

firms due to missing values in the Orbis data. The firms consist of 144 thousand establishments (including headquarters) and employ 6.4 million individuals. The data cover almost one third of total full time employment subject to social security contributions in Germany in 2012 (Bundesagentur für Arbeit, 2016).[9]

9 percent of firms are multi-establishment firms. They make up a disproportionate share of establishments and employment: 31 percent of establishments belong to and 34 percent of employees work for them. This pattern is similar across sectors. Multi-establishment firms account for 8 percent of firms, but 40 percent of employment in manufacturing, the sector with the highest number of firms. In retail and wholesale, the second largest sector, the share of multi-establishment firms is 12 percent, but their share in employment is 35 percent. On average across sectors, the share of multi-establishment firms in establishments and employment is three times their share in the number of firms.[10]

The statistics in the lower panel reflect the relevance of multi-establishment firms. Multi-establishment firms are substantially larger than single-estab-lishment firms in terms of their employment and sales. The median multi-establishment firm employs more than twice as many employees as the median single-establishment firm; at the 95th percentile, the factor is five. Median sales of multi-establishment firms are four times those of single-establishment firms.

Table 2 illustrates the complexity of multi-establishment firms. On average, multi-establishments firms have five establishments (including headquarters). Half of them have two establishments; the largest five percent have ten or more establishments (including headquarters). Multi-establishment firms are typically active in two sectors. The establishments tend to be geographically dispersed. Half of the multi-establishment firms have only establishments located within 170 km from their headquarters. At the top of the distribution, the distance between headquarters and establishments exceeds 540 km, about two thirds of the maximum possible distance within Germany. The distribution of the minimum area covered by firms with at least two establishments in addition to the headquarters is also skewed.

The lower panel of Table 2 illustrates differences between the headquarters and the other establishments. Headquarters are substantially larger than other establishments. The median headquarters is even larger than the median single-establishment firm. The size of the establishments varies with a larger standard deviation than the one for single establishment firms. This only partly reflects that the size cut-off is not binding at the establishment level. Multi-establishment firm management is concentrated at headquarters: Headquarters have a higher number of managerial layers, and a substantially higher management share than establishments.

[9]The total number of full time employees is only available for December 2012.

[10]The share of multi-establishment firms in the number of firms (employment) ranges from 5 (10) percent in construction (agriculture) to 17 (61) percent in mining and quarrying.

Table 2: 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
# sectors	9,812		1.6	0.9	1	2	3
Maximum distance to HQ, km	9,812		218	189	167	376	547
Minimum area covered, km^2	3,579		30,116	41,725	7,025	49,915	125,253
Descriptive statistics, est.	**N**	**HQ**	**Mean**	**SD**	**p50**	**p75**	**p95**
# employees	35,080	0	32	333	5	16	90
	9,812	1	107	669	27	76	365
# managerial layers	35,080	0	0.8	0.9	1	1	2
	9,812	1	1.7	1.1	2	3	3
Managerial share	35,080	0	22	28	9	39	81
(%, layers)	9,812	1	36	30	30	61	90
Managerial share	35,080	0	8	19	0	5	50
(%, Blossfeld)	9,812	1	10	16	4	14	41

Descriptive statistics, only ME firms. *# establishments (incl. HQ):* number of establishments (including headquarters); *# sectors:* number of three-digit sectors; *Maximum distance to HQ, km:* maximum distance between establishment and headquarters in kilometers; *Minimum area covered, km^2:* minimum area covered by establishments (including headquarters) in square-kilometers; *HQ:* indicator for headquarter establishment; *# managerial layers*: number of managerial layers, defined in section 2.2; *Managerial share (%, layers/Blossfeld):* share of wage sum earned by employees in managerial occupations (i.e., employees in managerial layers/managerial occupations according to Blossfeld (1983)); others: see Table 1.

3 Facts

This section describes the location and organization of multi-establishment firms. We first describe how geographic frictions between a location and the headquarters affect the location and size of the establishments. Taking the location as given, we then describe the managerial organization in the cross-section and over time.

3.1 Distance to headquarters decreases location probability

Table 3 describes the location pattern of multi-establishment firms. Columns 1 to 3 contain probit regressions that relate an indicator that is equal to one if a multi-establishment firm maintains an establishment in a county and county characteristics. Columns 4 to 6 contain OLS regressions that relate the log number of employees of an establishment to county characteristics. The regressions control for firm fixed effects to account for firm heterogeneity, so the sample only includes multi-establishment firms with at least two establishments.

Firms are the less likely to locate an establishment in a county, the more distant the county is from their headquarters. Establishment size also decreases with distance. A larger market potential increases the location probability. Lower wages and land prices in the county relative to the headquarters also relate positively to

Table 3: Location probability and establishment size, ME firms, 2012 cross section

Dependent variable	Location probability, probit			Log # est. employees, OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	-0.304^{***}	-0.293^{***}	-0.357^{***}	-0.106^{***}	-0.112^{***}	-0.137^{***}
	(0.021)	(0.023)	(0.020)	(0.018)	(0.019)	(0.017)
Log market potential	0.690^{***}	0.731^{***}		0.485^{***}	0.465^{***}	
	(0.026)	(0.030)		(0.044)	(0.046)	
Relative wages	-0.732^{***}	-0.667^{***}		-0.330^{**}	-0.433^{***}	
	(0.066)	(0.065)		(0.108)	(0.109)	
Relative land prices		-0.025^{***}			0.020^{***}	
		(0.005)			(0.005)	
# observations	3,719,275	3,225,429	3,719,275	21,496	19,203	21,496
# firms	9,275	8,741	9,275	3,006	2,773	3,006
Legal form FE	Y	Y	Y	N	N	N
HQ sector FE	Y	Y	Y	N	N	N
HQ county FE	Y	Y	Y	N	N	N
County FE	N	N	Y	N	N	Y
Firm FE	N	N	N	Y	Y	Y

The table presents the coefficients of a probit model (constant included; 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. ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1)-(3): indicator that is equal to one if ME firm i has an establishment in county c, (4)-(6): log number of employees of establishment. *Independent variables: Log distance to HQ:* log distance between county c and HQ county of firm i in km; *Log market potential:* distance weighted average of the GDP of county c and surrounding counties; *Relative wages/land prices:* wages/land prices in county c relative to wages in HQ county of firm i. We compute average wages in a county excluding firm i. Distance, market potential and relative land prices are based on data of the German Statistical Office. The number of observations varies due to covariate availability. FE abbreviates fixed effects.

the location probability. Similarly, larger market potential relates positively and higher wages relate negatively to establishment size. Unlike higher wages, higher land prices are positively related to establishment size.

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 relative wages indicate market-seeking and cost-cutting motives for having establishments. The different impact of land prices on the extensive and intensive margin is in line with considering land a fixed cost, so it is worthwhile to maintain only larger establishments at locations with higher land prices.

Fact 1 summarizes our findings:

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

Appendix B.1 shows that results are similar during the 2000-2010 period.

Table 4: Regression results, managerial organization of ME firms, 2012 cross-section

Dependent variable	# mgmt. layers, Poisson				Mg. share ∈ [0, 1], GLM			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.018*** (0.004)		0.019*** (0.004)		0.050*** (0.008)		0.029* (0.012)	
Log area		0.022*** (0.005)		0.026*** (0.004)		0.074*** (0.011)		0.074*** (0.014)
Log sales	0.125*** (0.004)	0.094*** (0.005)						
Log # non-mg. employees			0.139*** (0.004)	0.109*** (0.006)				
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
# firms	5,111	1,661	9,275	2,768	9,275	2,768	9,275	2,768

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 variable:* (1)-(4) number of managerial layers, (5),(6) managerial share in wage sum, defined by layer, (7),(8) managerial share in wage sum, defined by Blossfeld. *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. The number of observations is lower than the number of multi-establishment firms due to missing values for the legal form.

3.2 Distance to headquarters increases number of layers

Table 4 describes the relation of geographic frictions and firm organization as reflected by the number of managerial layers and the managerial share in the wage sum. For the number of managerial layers, we estimate Poisson regressions:

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

where 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 of the fractional nature of the managerial share, we follow Papke and Wooldridge (1996) and estimate a generalized linear model using the same covariates.[11] We approximate geographic frictions with the maximum distance to the headquarters of an establishment as well as the minimum area spanned by the establishments and the headquarters. The distance is defined for all multi-establishment firms, whereas the area is only defined for firms with at least two establishments. We use sales and the number of non-managerial employees as measures of firm size. Firm size controls for the positive effect of size on the number of layers (e.g., Caliendo et al., 2015b) and for the possibility of larger firms investing at farther destinations.

The regression results show that both distance and area have a positive impact

[11] We assume a logit link function and the binomial distributional family.

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

	Establishment			Headquarters		
Dependent variable	# layers	Mg. share ∈ [0, 1]		# layers	Mg. share ∈ [0, 1]	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.021**	0.026+	0.086**			
	(0.007)	(0.014)	(0.033)			
Maximum log distance to HQ				0.044***	0.109***	0.069***
				(0.004)	(0.009)	(0.014)
Log # non-mg. employees	0.309***			0.183***		
	(0.012)			(0.004)		
Model	Poisson	GLM	GLM	Poisson	GLM	GLM
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
# est./HQ	35,079	35,079	35,079	9,812	9,812	9,812

The table presents the coefficients. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1),(4) number of managerial layers, (2),(5) managerial share in wage sum, defined by layer, (3),(6) managerial share in wage sum, defined by Blossfeld. *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 subordinate establishment and headquarters in km; *Log # of non-mg. employees:* log number of employees at lowest layer in establishment/HQ.

on the number of managerial layers in a firm. According to columns 1 and 3, doubling the maximum distance of an establishment to the headquarters is associated with the same increase of the number of layers as 14 percent higher sales or more non-managerial employees.The impact of the area in columns 2 and 4 is similar.[12] Doubling the area is associated with the same change of layers as increasing size by 25 percent. Likewise, the managerial share is positively related both to the maximum distance of the establishments and the area they span.

The firm-level results may disguise heterogeneity at the establishment level. As the descriptive statistics in Table 2 indicate, the managerial organization of the establishments does not copy the headquarters. 73 percent of establishments have fewer managerial layers than the headquarters. Table 5 complements the firm-level estimates with establishment and headquarter-level regressions. Columns 1 to 3 refer to the establishments and columns 4 to 6 refer to the headquarters.

Similar to the firm-level results, the number of managerial layers and the managerial share of an establishment increase if the establishment is located at a larger distance from the headquarters. The number of managerial layers and the managerial share at the headquarters also respond positively to geographic frictions. Establishment size correlates positively with the number of layers both for headquarters and establishments.

[12] The number of observations is lower than the total number of firms with three establishments, because the area cannot be computed if the firm has several establishments in the same county.

Fact 2 summarizes our findings:

Fact 2. *The number of managerial layers and the managerial share of multi-establishment firms correlate positively with the distance between the headquarters and the establishments and the area that they span, conditional on other firm characteristics. The managerial organization of both the establishments and the headquarters responds to geographic frictions.*

Robustness. Appendix section B.2 documents the robustness of our results. Tables B.2 and B.3 replicate the regression results in Tables 4 and 5 using linear models. Tables B.4 to B.6 explore whether distance takes up non-linear effects of firm and establishment size on firm organization. Tables B.4 and B.5 include squared firm and establishment size as covariates. Table B.6 includes the number of establishments of a firm as covariate in the firm level regressions. The results are robust to all of these changes. Tables B.7 and B.8 explore possible non-linear effects of geography on the managerial organization by including quartile dummies of the (maximum) distance to headquarters. We find the strongest effects for the third and forth distance quartile. Multi-establishment firms may separate management and production geographically. Table B.9 replicates the firm-level analysis treating the lowest-level layer in each firm as non-managerial layer. Finally, Table B.10 presents 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. Results are very similar for "GmbHs" and "GmbH & Co. KGs", the two most popular legal forms. Coefficients are insignificant for AGs, which is likely due to the fact that there are only few AGs in the sample.[13]

3.3 Multi-establishment firms reorganize gradually

Table 6 presents the reorganization pattern of multi-establishment firms over time. The upper panel displays the percentage share of firms that transition from a number of managerial layers in year t to a possibly different number of managerial layers in year $t+1$. The managerial organization is sluggish: at least 80 percent of firms keep constant the number of managerial layers across periods. Firms that change the number of layers add or drop one layer. The upper panel of Appendix Table B.12 shows that changes of the number of layers are related to changes of firm size. These patterns are similar to the reorganization patterns reported for French and Danish firms in the literature (Caliendo et al., 2015b, Friedrich, 2016) and to the patterns of single-establishment firms reported in Appendix Table B.11.

[13] A "GmbH" is a limited liability company. A "GmbH & Co. KGs" is a limited partnership with a limited liability company as general partner. "AGs" are public companies.

Table 6: Transition dynamics of the managerial organization, 2000-2010 panel

(a) # managerial layers, firm-level

# 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 possibly 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 structure in year t (given in the rows) to a possibly different managerial structure 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.

To understand whether multi-establishment firms change their organization at the same time or at different points in time at different establishments, the lower panel displays the organizational dynamics at the level of the headquarters and establishments. We count the number of managerial layers at the headquarters and the maximum number of managerial layers at the establishment to account for a possibly different number of establishments across firms.

The managerial organization at the establishment level is less stable than the managerial organization at the firm level: there is less mass on the diagonal of the lower panel than on the diagonal of the upper panel. Notably, we find that multi-establishment firms reorganize gradually. They add or drop layers at either the headquarters or the establishment(s). For example, among multi-establishment firms with two layers both at the headquarters and the establishments, 9 percent add a layer at the headquarters and 10 percent drop a layer at the establishments. The latter adjustment does not show up as reorganization at the firm level. Only 1 percent of firms choose a lower or higher number of layers across both types of establishments. Overall, among the firms that reorganize, 44 percent change the number of layers only at the headquarters, 47 percent change it only at the establishments, and only 9 percent change it at both types of establishments.[14] The lower panel of Appendix Table B.12 shows that changes to firm organization are related to changes of firm size also at the establishment and headquarter-level.

Fact 3 summarizes our finding.

Fact 3. *Multi-establishment firms reorganize gradually as they grow: they typically add or drop layers either at the headquarters or at the establishments.*

Robustness. Table 6 aggregates across the establishments of a firm. To make sure that the aggregation does not yield misleading results, Appendix Table B.13 shows that the results are similar for firms with headquarters and exactly one establishment and firms with headquarters and at least two establishments. To explore whether geography affects the transition patterns, Appendix Table B.14 displays results for firms with only proximate and firms with distant establishments. We split the sample at the median of the maximum log distance of establishments from the headquarters (170 km). The transition patterns are similar across the two groups.

4 Model

To explain why the location of establishments affects the managerial organization, we develop a model where firms endogenously choose whether to operate an establishment and the managerial organization.

4.1 Set-up

We consider an economy with two locations, $j = \{0, 1\}$. The N_j agents per location each supply one unit of time to the labor market. The agents are immobile, so local wages w_j can differ. The agents derive utility from consuming differentiated products i:

$$U(x(\alpha_i)) = \left(\int_A \alpha_i^{\frac{1}{\sigma}} x(\alpha_i)^{\frac{\sigma-1}{\sigma}} dG(\alpha) \right)^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

$x(\alpha_i)$ is an agent's consumption of product i, $\alpha_i > 0$ is the agents' taste for product i, A is the set of all available products, and $\sigma > 1$ is the elasticity of substitution. The taste draws α_i follow the distribution $G(\alpha)$. Each firm makes exactly one product, so we use the index i interchangeably firms and products.

Production. Production is a problem solving process based on labor and knowledge (as in Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). Every unit of labor

[14] These 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 not included in Table 6.

employed in production generates a unit mass of problems. Problems are production possibilities: the labor input turns into output if the problems are solved using knowledge. Mathematically, knowledge is an interval ranging from zero to an upper bound. We denote the length of a knowledge interval by z. A problem is solved if it is realized within the knowledge 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. 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 on the labor market to supply labor and knowledge for production. The firm's employees supply labor by spending their time generating problems. To use knowledge in production, the employees have to learn it. They spend $w_j c z$ to learn a knowledge interval of length z, where c denotes the learning cost that is equal across locations. As is standard in the literature (e.g., Caliendo and Rossi-Hansberg, 2012), the firm remunerates the employees for their time and their learning expenses, so employees receive remuneration $w_j(1 + cz)$.

The employees of the firm can communicate problems with each other, so they can leverage differences in their knowledge. Communication is costly: an employee in location j spends θ_{kj} units of time helping an employee in location k. Helping an employee in another location is more costly than helping in the same location: $1 > \theta_{10} \geq \theta_{00} > 0$. 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 has to find a competent fellow employee.

Organization. Firms organize their employees in hierarchical layers (as in Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). We call the employees at the lowest layer $\ell = 0$ production workers. They supply labor to generate problems and solve those that are realized in their knowledge interval. We call the employees at the higher layers $\ell \geq 1$ managers. They supply only knowledge for production and spend their time helping the employees at the next lower layer.[15] The highest managerial layer consists of the CEO. All firms consist of production workers and a CEO; they may have one or more layers of below-CEO managers. The knowledge levels of the employees are overlapping, so employees at layer ℓ know the knowledge of employees

[15]The optimal knowledge hierarchy features specialization and organization by frequency, i.e. only the lowest layer supplies labor and the knowledge of higher layers solves the rarer problems (Garicano, 2000).

at layer $\ell - 1$ and more.[16] Consequently, CEO knowledge $\bar{z}$ delimits the maximum possible output per unit of labor input, because the CEO is the most knowledgeable employee of the firm. An important assumption is that each firm has exactly one CEO. The CEO is therefore a resource of limited supply for a firm.

The helping costs θ_{jk}, the learning costs c, the predictability of the production process λ and the tastes α_i are exogenous parameters. Assumption 1 in the Appendix restricts the possible parameter values. The model is partial equilibrium, so the wages w_j are also given. To simplify the exposition, sections 4.2 and 4.3 study the organization of a firm in location 0 taking as given output $\tilde{q}$. Section 4.4 endogenizes output based on the competition among many firms i in the product market.

4.2 Single-establishment firm organization

We first determine the optimal organization of a firm that only produces in the headquarters.[17] The organization consists of the number of below-CEO managerial layers $L \in \{0, 1, 2\}$, the number $n_{0,L}^{\ell}$ and knowledge level $z_{0,L}^{\ell}$ of employees per layer $\ell = 0, ..., L$, and the knowledge of the CEO $\bar{z}_{0,L}$.[18] The indexes $0, L$ refer to the location of the firm $j = 0$ and the number of below-CEO managerial layers L and reflect that these variables affect the other choices.

The optimal number of below-CEO layers yields minimal production costs:

$$C(\tilde{q}) = \min_{L \in \{0,1,2\}} \tilde{C}_{0,L}(\tilde{q}). \tag{2}$$

The optimal number and knowledge levels of employees minimize costs for a given number of layers:

$$C_{0,L}(\tilde{q}) = \min_{\{n_{0,L}^{\ell}, z_{0,L}^{\ell}\}_{\ell=0}^{L}, \bar{z}_{0,L} \geq 0} \sum_{\ell=0}^{L} n_{0,L}^{\ell} w_0 \left(1 + c z_{0,L}^{\ell}\right) + w_0 \left(1 + c \bar{z}_{0,L}\right) \tag{3}$$

$$\text{s.t.} \quad n_{0,L}^{0} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right) \geq \tilde{q} \tag{4}$$

$$1 \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{L}} \tag{5}$$

$$n_{0,L}^{\ell} \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{\ell-1}} \quad \forall \ell = 1, ..., L \tag{6}$$

$$\bar{z}_{0,L} \geq z_{0,L}^{L}, \quad z_{0,L}^{\ell} \geq z_{0,L}^{\ell-1} \quad \forall \ell = 1, ..., L \tag{7}$$

The production costs consist of the personnel costs for the employees and the CEO. Constraint (4) specifies that the number of production workers and CEO knowledge have to suffice to produce output $\tilde{q}$. Constraints (5) and (6) reflect that the amount

[16] Overlapping knowledge levels simplify the analysis as overlaps and gaps between CEO and establishment knowledge may occur with non-overlapping knowledge and multiple establishments.

[17] Our results conform with Caliendo and Rossi-Hansberg (2012)'s for non-overlapping knowledge.

[18] We restrict our attention to $L + 1 \leq 3$ managerial layers in line with sections 2 and 3.

of time of the CEO and the managers limit the number of problems that can be communicated to them. This number is computed as the number of problems, $n^0_{0,L}$, multiplied with the helping costs, θ_{00}, and the probability that the problem is not yet solved, $e^{-\lambda z^{\ell-1}_{0,L}}$. Finally, knowledge levels are overlapping and positive (constraint 7).

Appendix C.1.1 contains the Lagrangian equation and the first order conditions. Two multipliers from the Lagrangian equation help characterizing the organization. The multiplier for constraint (4), $\xi_{0,L}$, denotes the marginal production costs. The multiplier for constraint (5), $\varphi_{0,L}$, denotes the marginal benefit of CEO time. CEO time is fixed at one unit. $\varphi_{0,L}$ reflects how costly this constraint is for the firm.

The first order conditions show that the firm optimally chooses CEO knowledge such that its marginal benefit and its marginal cost are equal:

$$w_0 c = \lambda e^{-\lambda \bar{z}_{0,L}} \xi_{0,L} n^0_{0,L} \tag{8}$$

The marginal cost of CEO knowledge consists of the increase of CEO remuneration $w_0 c$. The marginal benefit is the reduction of production costs, because more output is producible for every unit of labor input with higher CEO knowledge.

Given CEO knowledge, constraint (4) determines the number of production workers $n^0_{0,L}$. Constraint (5) determines the knowledge level of the highest below-CEO layer in the firm. The employees at the highest below-CEO layer have to solve a sufficient fraction of problems so only the unit of CEO time is used. The knowledge levels of the production workers and managers at lower layers are a recursive function of the knowledge level at the highest below-CEO layer:

$$e^{\lambda\left(z^{\ell-1}_{0,L} - z^{\ell-2}_{0,L}\right)} = \left(1 + c z^{\ell}_{0,L}\right) \frac{\lambda}{c} \quad \forall \ell = 2, ..., L, \tag{9}$$

$$e^{\lambda z^0_{0,L}} = \left(1 + c z^1_{0,L}\right) \frac{\lambda \theta_{00}}{c}. \tag{10}$$

At each layer, the firm trades off the costs of higher knowledge in terms of higher remuneration and the benefit of a lower number of employees at the next higher layer. Constraint (6) determines the number of middle managers. Finally, the marginal production costs $\xi_{0,L}$ and the marginal benefit of CEO time $\varphi_{0,L}$ are given by:

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

$$\varphi_{0,L} = \frac{w_0 c}{\lambda} e^{\lambda\left(z^{L-1}_{0,L} - z^{L-2}_{0,L}\right)} \quad \text{for } L \geq 1, \quad \varphi_{0,L} = \frac{w_0 c}{\lambda \theta_{00}} e^{\lambda z^0_{0,L}} \quad \text{for } L = 0.$$

Understanding how output $\tilde{q}$ affects firm organization is useful for the later analysis of multi-establishment firms.

Proposition 1. *Given the number of below-CEO managerial layers L of the firm,* *a) 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 knowledge of the CEO $\bar{z}_{0,L}$ and the marginal benefit of CEO time $\varphi_{0,L}$ increase with output $\tilde{q}$.*

b) The cost function $C_{0,L}(\tilde{q})$ strictly increases with output $\tilde{q}$. The average cost function $AC_{0,L}(\tilde{q})$ is U-shaped. It reaches a minimum at $\tilde{q}^{*}_{L}$ where it intersects the marginal cost function, and converges to infinity for $\tilde{q} \to 0$ and $\tilde{q} \to \infty$.

Proof. See Appendix C.1.2. □

CEO knowledge $\bar{z}_{0,L}$ and the number of production workers $n^{0}_{0,L}$ increase because labor and knowledge are complementary inputs in production, so the firm optimally employs a higher amount of both to achieve higher output. The larger the output is, the more problems the production workers generate and, if unsolved, communicate to higher layers. Higher output therefore increases the number of employees $n^{\ell}_{0,L}$ at all below-CEO layers. The amount of CEO time is fixed, however. The knowledge of the employees at the highest below-CEO layer $z^{L-1}_{0,L}$ increases, because otherwise, the CEO could not help with all problems that are communicated to him. The knowledge levels at lower layers $z^{\ell}_{0,L}$, $\ell = 0, ..., L-1$ also increase, though to a lesser extent, thereby mitigating the increase in the number of employees at the below-CEO layers. The larger the firm is, the more beneficial it would be to increase CEO time and avoid the increase of knowledge. Thus, the shadow price of the CEO time constraint—the marginal benefit of CEO time—increases with output.

The resulting cost function is strictly increasing, as the marginal costs are positive. The average cost function is U-shaped. The U-shape reflects two counteracting forces. On the one hand, the marginal cost of production increase with output.[19] On the other hand, the quasi-fixed costs of the CEO and the middle managers are spread over a larger output. For quantities below the minimum efficient scale, $\tilde{q} < \tilde{q}^{*}_{L}$, the latter effect dominates; for quantities above, $\tilde{q} > \tilde{q}^{*}_{L}$, the former effect dominates.

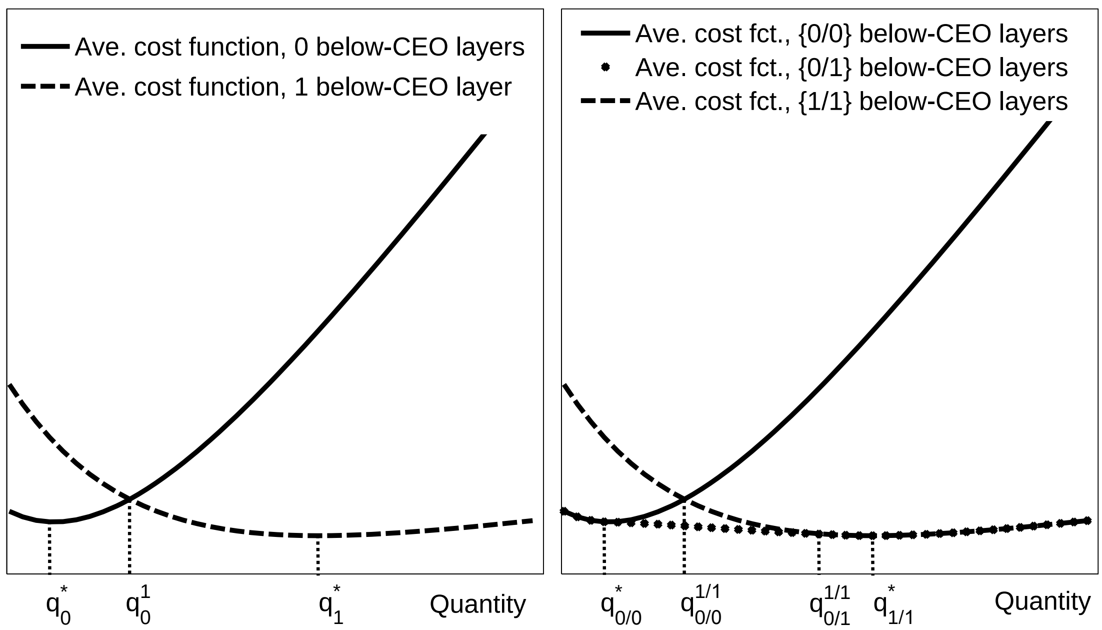
The optimal number of managerial layers minimizes the production costs. The firm faces a trade-off. On the one hand, middle managers entail a quasi-fixed cost, because they are remunerated, but do not generate problems and increase output. On the other hand, middle managers solve part of the problems that are generated by the production workers and reduce the number of problems sent to the CEO. They thus allow decreasing the knowledge of the production workers and the marginal production costs. In consequence, adding a layers is only worthwhile if the firm is sufficiently large. The optimal number of layers increases with output $\tilde{q}$.

Figure 1a illustrates the choice of adding a layer of middle managers using the average cost function of a firm with only a CEO ($L = 0$) or a CEO and middle

[19] The marginal cost of production globally increase with output for $L = 0$. For $L > 0$, they increase for sufficiently high output; in particular, they increase at the minimum efficient scale.

Figure 1: Illustration of the average cost functions

(a) Single-establishment firm (b) Multi-establishment firm



The figure plots the average cost functions of a single and multi-establishment firm for $w_0 = w_1$, $\theta_{00} = \theta_{10}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$. (a): The average cost function of a single-establishment firm is U-shaped for a given number of below-CEO layers. The firm adds a layer at the intersection $\tilde{q}_0^1$. (b): The average cost function of a multi-establishment firm with a symmetric number of below-CEO layers $\{0, 0\}$ or $\{1, 1\}$ is U-shaped. The firm adds a layer at one establishment at the minimum efficient scale $\tilde{q}_0^*$ and a layer at the other establishment at $\tilde{q}_{0/1}^{1/1} > \tilde{q}_0^1$.

managers ($L = 1$). The minimum efficient scale $\tilde{q}_L^*$ of an organization increases with the number of below-CEO layers, reflecting the higher quasi-fixed costs of more managers. The firm adds a layer at the crossing $\tilde{q}_0^1$ (see also Appendix C.1.3).

4.3 Multi-establishment firm organization

The firm may maintain an establishment at location $j = 1$ to exploit wage differences or to access the local product market. To distinguish the effect of wage differences and market access on the organization, we derive the optimal multi-establishment firm organization in two steps. In section 4.3.1, the firm can hire employees in the local labor markets at both locations, but sells output in a global product market. We assume that $w_0 \geq w_1$. In section 4.3.2, the firm incurs transport costs to ship output from one location to the other.

4.3.1 Local labor markets, global product market

The CEO is located in 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 managerial layers L_j per location. We index the number of below-

CEO layers by j for clarity. We use the term "organizational structure" and the variable ω to denote the combination of the number of below-CEO layers. All other endogenous variables depend on the location and the organizational structure, so we index them by j, ω. If the firm maintains an establishment, it chooses how much output $q_{j,\omega}$ and which share $s_{j,\omega}$ of CEO time to allocate to the headquarters and the establishment. The firm also determines the level of CEO knowledge $\bar{z}_{0,\omega}$ as well as the number $n_{j,\omega}^{\ell}$ and knowledge level $z_{j,\omega}^{\ell}$ of the employees in each layer ℓ.

We split the optimization problem into three steps. First, the firm chooses the optimal organizational structure ω from the set of possible structures Ω to minimize its production costs, similarly to choosing the number of layers in section 4.2:

$$C\left(\tilde{q}\right) = \min_{\omega \in \Omega} \tilde{C}_{0,\omega}\left(\tilde{q}\right) \tag{11}$$

Second, the firm determines the production quantities $q_{j,\omega}$ and the allocation of CEO time $s_{j,\omega}$ and chooses CEO knowledge $\bar{z}_{0,\omega}$ to minimize the costs of the chosen organizational structure. The costs consist of the personnel costs per establishment and the remuneration of the CEO time that is not used in production.

$$\tilde{C}_{0,\omega}\left(\tilde{q}\right) = \min_{\{q_{j,\omega}, s_{j,\omega}\}_{j=0}^{1}, \bar{z}_{0,\omega} \geq 0} \sum_{j=0}^{1} C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0 \left(1 + c\bar{z}_{0,\omega}\right) \tag{12}$$

$$\text{s.t.} \quad s_{0,\omega} + s_{1,\omega} \leq 1 \tag{13}$$

$$q_{0,\omega} + q_{1,\omega} \geq \tilde{q} \tag{14}$$

Equation (13) reflects the CEO's time constraint. Constraint (14) describes that the production quantities have to sum up at least to the total output $\tilde{q}$.

Third, the firm determines the number of employees and their knowledge for each below-CEO layer. If the firm decides to produce at a location, the production costs consist of the below-CEO personnel costs as well as the remuneration for the CEO time allocated to it. Otherwise, the production costs are zero.

$$C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) \begin{cases} \overset{q_{j,\omega}>0}{=} \min\limits_{\{n_{j,\omega}^{\ell}, z_{j,\omega}^{\ell}\}_{\ell=0}^{L_j} \geq 0} \sum\limits_{\ell=0}^{L_j} n_{j,\omega}^{\ell} w_j \left(1 + c z_{j,\omega}^{\ell}\right) + s_{j,\omega} w_0 \left(1 + c\bar{z}_{0,\omega}\right) \\ \overset{q_{j,\omega}=0}{=} 0 \end{cases} \tag{15}$$

$$\text{s.t.} \quad n_{j,\omega}^{0}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right) \geq q_{j,\omega} \tag{16}$$

$$s_{j,\omega} \geq n_{j,\omega}^{0} \theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} \tag{17}$$

$$n_{j,\omega}^{\ell} \geq n_{j,\omega}^{0} \theta_{jj} e^{-\lambda z_{j,\omega}^{\ell-1}} \quad \forall \ell = 1, ..., L_j \tag{18}$$

$$\bar{z}_{0,\omega} \geq z_{j,\omega}^{L_j}, \quad z_{j,\omega}^{\ell} \geq z_{j,\omega}^{\ell-1} \quad \forall \ell = 1, ..., L_j \tag{19}$$

The constraints (16)-(19) are analogous to the constraints (4)-(7).

We solve the problem backwards. We first determine the number and knowledge of employees per layer, taking as given the firm level choices as well as the organizational structure. We then solve for the knowledge of the CEO, the allocation of CEO time and the production quantities given the organizational structure, which we determine in the last step. Appendix C.2.1 contains the Lagrangian equations and the first order conditions.

Establishment-level choices. The establishment outcomes depend on the choices at the firm level—CEO knowledge, the production quantities and the allocation CEO time—through the binding constraints (16)-(18). The formal expressions are variants of those in section 4.2, which is why we state them in Appendix C.2.1.

Constraint (16) determines the number of production workers. Constraint (17) fixes the knowledge level of the highest below-CEO layer at the headquarters or establishment as a function of the allocated share of CEO time. The knowledge levels at lower layers are recursive functions of the knowledge level at the highest below-CEO layer. Constraint (18) determines the number of middle managers. The Lagrangian multipliers $\xi_{j,\omega}$ denote the marginal production costs and the multipliers $\varphi_{j,\omega}$ denote the marginal benefit of CEO time at a location.

Firm-level choices. The firm balances the marginal benefit and marginal cost of CEO knowledge, analogously to section 4.2:

$$w_0 c = \lambda e^{-\lambda \bar{z}_{0,\omega}} \sum_{j=0}^{1} \xi_{j,\omega} n_{j,\omega}^{0}. \tag{20}$$

The firm optimally uses the full unit of CEO time and produces only the given output, i.e. the constraints (13) and (14) are binding. The firm can reduce the production costs by reallocating the production quantity or CEO time as long as the marginal production costs or the marginal benefit of CEO time are not equal at the headquarters and the establishment.

Proposition 2. *Suppose the firm produces in the headquarters and the establishment. The firm allocates output to equalize the marginal production costs and CEO time to equalize the marginal benefit of CEO time at the headquarters and the establishment. Formally, in optimum,*

$$\xi_{0,\omega} = \xi_{1,\omega} \quad \text{and} \tag{21}$$

$$\varphi_{0,\omega} = \varphi_{1,\omega}. \tag{22}$$

Proof. See Appendix C.2.2. □

Hence, the firm produces total output at one location if the marginal costs of total output at this location are lower than the marginal costs at the other location. It spends the full unit of CEO time for one location if the marginal benefit of doing so exceeds the marginal benefit of spending time for the other location.

Corollary 1. *It is not optimal to produce at an establishment with the same number of below-CEO management layers as the headquarters $L_0 = L_1$ if the helping costs across space exceed those within a location, $\theta_{10} > \theta_{00}$, but the wages are equal or higher at the establishment than at the headquarters, $w_1 \geq w_0$.*

Proof. See Appendix C.2.2. □

Intuitively, the firm only produces in the establishment if some advantage there counterbalances the higher helping costs across space θ_{10}. The advantage can consist of lower wages or a different managerial structure of the establishment.

Comparative statics. To derive the optimal organizational structure ω, it is useful to understand how firm choices depend on the output $\tilde{q}$ and the helping costs θ_{10}.

Proposition 3. *Suppose the firm produces in the headquarters and the establishment. Suppose further that the headquarters and the establishment are asymmetric, 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 $\bar{z}_{0,\omega}$ and the total number of employees at all below-CEO layers $\sum_{j=0}^{1} n_{j,\omega}^{\ell}$, $\forall \ell < L$, increase with output $\tilde{q}$, while the knowledge of the employees at all below-CEO layers $z_{j,\omega}^{\ell}$ and the marginal benefit of CEO time $\varphi_{j,\omega}$ are constant.

b) The share of CEO time $s_{j,\omega}$ and the number of employees at all below-CEO layers $n_{j,\omega}^{\ell}$, $\forall \ell < L$, increase with output $\tilde{q}$ at the establishment and decrease at the headquarters, unless $L_0 > L_1$ and wages w_1 are too high. The production quantity $q_{1,\omega}$ increases at the establishment.

c) The cost function $C_{0,\omega}(\tilde{q})$ strictly increases with output $\tilde{q}$. The marginal production costs $\xi_{j,\omega}$ decrease with output $\tilde{q}$.

Under symmetry, i.e., $\theta_{10} = \theta_{00}$, $w_1 = w_0$ and $L_1 = L_0$, output has the same effect on the choices of a multi-establishment firm as in Proposition 1.

Proof. See Appendix C.2.3. □

As in Proposition 1, higher total output $\tilde{q}$ leads to higher CEO knowledge and a higher total number of production workers because labor and knowledge are complementary inputs. The higher number of production workers leads to a higher number of employees at all below-CEO layers. If there is asymmetry between the establishment and the headquarters, the below-CEO knowledge levels and the marginal benefit of CEO time do not vary with output. This is because maintaining an establishment different from the headquarters effectively allows the firm to use labor and knowledge in two different ways, and to increase output by recombining them. For instance, the number of layers may differ. Section 4.2 shows that the efficiency of a certain number of layers depends on output for a single-establishment firm. The multi-establishment firm can freely allocate output between the headquarters and the establishment and thus choose the optimal combination of layers for its output. Similarly, wages may differ between locations. The multi-establishment firm can allocate output to optimally combine the lower helping costs at the headquarters and the lower wages at the establishment. The firm therefore does not vary the knowledge levels with output, but chooses the knowledge levels that are optimal given location characteristics. It accommodates higher output by reallocating CEO time and production quantities between locations. If the firm instead produces either only at the headquarters or the establishment, it increases below-CEO knowledge to respect the CEO's time constraint (i.e., Proposition 1 applies).

The production quantities and the allocation of CEO time reflect that the firm leverages the asymmetries between the headquarters and the establishment. The larger the firm is, the more important are low wages relative to low helping costs, because the firm hires more employees. The firm thus allocates higher shares of output and CEO time to the establishment. The number of employees at a location depends on the share of CEO time and thus varies with it.

The marginal production costs are positive, so the cost function increases with output. The marginal production costs decrease with output. This property results because below-CEO knowledge levels are constant, so the costs per unit of labor input are constant. CEO knowledge increases with output. Therefore, more output is producible for every unit of labor input, which decreases the marginal costs.

If the headquarters and the establishment are fully symmetric with respect to both location characteristics and the number of below-CEO layers, the multi-establishment firm makes the same choices as if it produced only at the headquarters. Consequently, changes in output affect multi-establishment firm organization as stated in Proposition 1.

The helping costs θ_{10} also affect the firm choices.

Proposition 4. *Suppose the firm produces in the headquarters and the establishment. Suppose further that $\theta_{10} > \theta_{00}$ and $L_0 \leq L_1$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$, the marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\forall \ell < L$, increase with the helping *costs θ_{10}. The increase is stronger at higher than at lower layers. If the number of below-CEO layers is equal across locations, the increase is stronger at the establishment than at the headquarters. The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n^{\ell}_{j,\omega}$, $\forall \ell < L$, decreases.*

b) The share of CEO time $s_{j,\omega}$, the production quantity $q_{j,\omega}$ and the number of employees at all below-CEO layers $n^{\ell}_{j,\omega}$ increase with the helping costs θ_{10} at the headquarters and decrease at the establishment.

c) The cost function $C_{0,\omega}(\tilde{q})$ and the marginal production cost $\xi_{j,\omega}$ increase with the helping costs θ_{10}.

The comparative statics also hold for a higher number of below-CEO layers at the headquarters than at the establishment, $L_0 > L_1$, if wages w_1 are sufficiently low.

Proof. See Appendix C.2.4. □

Higher helping costs θ_{10} make it more costly to use CEO knowledge because the CEO spends more time per problem. The firm increases CEO knowledge to compensate the higher costs with a higher benefit of using the CEO. Due to the CEO time constraint, more problems have to be solved at below-CEO layers. The marginal benefit (or shadow price) of CEO time and the knowledge at the below-CEO layers increase. The increase is stronger at higher than at lower layers because the number of employees is lower at higher layers, so it is cheaper to increase their knowledge. Below-CEO knowledge increases at the establishment and the headquarters, because the marginal product of knowledge is decreasing. Consequently, fully compensating the higher helping costs in the establishment is not efficient. The knowledge increase is stronger at the establishment than at the headquarters to ensure that the marginal benefit of CEO time remains equal. The higher knowledge levels allow the firm to hire fewer production workers and fewer managers at all below-CEO layers.

Higher helping costs θ_{10} make it more costly to produce at location 1. Correspondingly, the firm reduces the share of CEO time, the production quantity, and the number of employees there. As the organizational adjustments do not fully compensate the higher helping costs, the cost function and the marginal costs increase.

An important implication of Proposition 4 is that the organization of the multi-establishment firm is interdependent across the headquarters and the establishment. Changes of the helping costs at the establishment lead to organizational adjustments at the headquarters and the establishment because of the shared CEO. This interdependence is also reflected in the choice of organizational structure.

Organizational structure. The firm chooses the organizational structure with the minimal production costs. To disentangle the impact of total output and of location characteristics on the organization, we first derive the optimal organization when both wages and helping costs are equal, $w_0 = w_1$ and $\theta_{00} = \theta_{10}$.

Proposition 5. *Suppose that wages and helping costs are equal, $w_0 = w_1, \theta_{00} = \theta_{10}$. Let "$\{L_0/L_0\}$-organization" denote the organizational structure of a multi-establishment firm with L_0 below-CEO layers at the headquarters and the establishment, i.e. $\omega = L_0/L_0$. Let "$\{L_0/L_0+1\}$-organization" denote the organizational structure of a multi-establishment firm with L_0 below-CEO layers at the headquarters and L_0+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 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 $\tilde{q}^*_{L_0+1/L_0+1} > \tilde{q} > \tilde{q}^*_{L_0/L_0}$.

c) The average cost function of the $\{L_0+1/L_0+1\}$-organization intersects the average cost function of the $\{L_0/L_0\}$-organization at the output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$ between the minimum efficient scales, i.e., $q^*_{L_0+1/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0} > q^*_{L_0/L_0}$. The average cost function of the $\{L_0/L_0+1\}$-organization intersects the average cost function of the $\{L_0+1/L_0+1\}$-organization at the higher output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$.

As a result, the multi-establishment firm with L_0 below-CEO layers at the headquarters and the establishment reorganizes gradually with higher output: it adds a layer of managers at either location at the output $\tilde{q}^*_{L_0/L_0}$ and at the other location at output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0+1} \in (\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}, \tilde{q}^*_{L_0+1/L_0+1})$.

Proof. See Appendix C.2.6. □

Figure 1b illustrates the average costs of the multi-establishment firm, taking an organization with 0 or 1 below-CEO layers as example. The figure shows that the average cost function of an organization with the same number of below-CEO layers at the headquarters and the establishment is U-shaped (Proposition 5a)). It coincides with the average cost function of a single-establishment firm. The average cost function of the $\{0/0\}$-organization increases for quantities above the minimum efficient scale $q^*_{0/0}$. In contrast, the average cost function of the $\{0/1\}$-organization decreases (part b)). Consequently, the former intersects the average cost function of the $\{1/1\}$-organization at a lower quantity than the latter (part c)).[20]

[20]The average cost function of the $\{0/1\}$-organization coincides with the average cost functions of the 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.

Proposition 5 is a key result of the model. It states that the multi-establishment firm reorganizes gradually as it grows. If the firm produces at either the headquarters or the establishment, hiring a layer of middle managers is only worthwhile for sufficiently high output. The multi-establishment firm is free to allocate output and CEO time. It can optimally combine different numbers of below-CEO layers at the headquarters and the establishment and thus decrease production costs. At the quantity $q^*_{L_0/L_0}$, the $\{L_0/L_0\}$-organization has the minimum average costs. A multi-establishment firm with a $\{L_0/L_0+1\}$-organization would allocate total output to the headquarters with L_0 below-CEO layers at $q^*_{L_0/L_0}$. For higher output $\tilde{q} > q^*_{L_0/L_0}$, the average costs of the $\{L_0/L_0\}$-organization increase, because it exceeds its minimum efficient scale. 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 can allocate a share of output to the establishment with L_0+1 below-CEO layers. For output close to the minimum efficient scale, only a small share is allocated to the establishment, but the larger the output $\tilde{q}$, the larger its share of production.

The additional managerial layer at the establishment releases CEO time: relative to output, the CEO spends a larger share of time at the headquarters with L_0 than at the establishment with L_0+1 below-CEO layers. This keeps below-CEO knowledge low. The additional managerial layer thus increases efficiency both at the establishment and the headquarters. In consequence, the multi-establishment firm only switches to the $\{L_0+1/L_0+1\}$-organization at output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$. As in Proposition 4, the organization of a multi-establishment firm is interdependent: The optimal number of layers at the headquarters depends on the number of layers at the establishment.

Possible asymmetries of wages and helping costs qualify, but do not fundamentally alter the results.

Corollary 2. *Suppose the helping costs across space exceed those within a location, $\theta_{10} > \theta_{00}$, but the wages are equal or higher at the establishment than at the headquarters, $w_1 \geq w_0$. As total output $\tilde{q}$ increases, the firm alternates between production only at the headquarters and multi-establishment production with an unequal number of below-CEO layers. The higher the helping costs θ_{10} are, the lower is the range of output levels for which multi-establishment production is optimal.*

Proof. See Appendix C.2.6. □

Corollary 2 is an immediate implication of Corollary 1: multi-establishment production is only optimal if the higher helping cost across space are counterbalanced by lower wages or a lower-cost organization at the establishment. The firm thus only chooses multi-establishment production with an unequal number of below-CEO layers. The higher the helping costs across space are, the smaller is the range of output levels for which the organizational advantage can counterbalance them. In contrast, if lower wages at the establishment $w_1 < w_0$ counterbalance the higher helping costs across space, multi-establishment production can be optimal both with the same and a different number of below-CEO layers at the establishment.

4.3.2 Local labor and local product markets

The firm may maintain the establishment not only to leverage wage differences across locations, but also to be closer to its customers. To capture market access motives, we assume that the firm incurs transport costs if it sells output produced at one location in the product market at the other location. We assume that transport costs $\tau > 1$ are iceberg-type.[21] We take as given the potentially different amounts of output $\{\tilde{q}_j\}_{j=0}^1$ that the firm supplies to each market.

The transport costs do not affect the choice of the optimal organizational structure (equation 11) or the optimization problem at the level of the headquarters and the establishment (equations 15-19). However, they constrain the choice of the production quantities at the firm level.

$$\tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^1\right) = \min_{\{q_{j,\omega},s_{j,\omega}\}_{j=0}^1,\bar{z}_{0,\omega}\geq 0} \sum_{j=0}^{1} C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0 \left(1 + c\bar{z}_{0,\omega}\right) \quad (23)$$

$$\text{s.t.} \quad s_{0,\omega} + s_{1,\omega} \leq 1 \quad (24)$$

$$\mathbb{1}(q_{j,\omega} \geq \tilde{q}_j \wedge q_{k,\omega} \leq \tilde{q}_k)(q_{j,\omega} - \tilde{q}_j + \tau(q_{k,\omega} - \tilde{q}_k)) \geq 0, \quad k \neq j \quad (25)$$

Constraint (25) states that the production quantity at location j, $q_{j,\omega}$, has to cover local output $\tilde{q}_j$ and a possible difference between local production and local output at location k including transport costs: $q_{j,\omega} \geq \tilde{q}_j + \tau(\tilde{q}_k - q_{k,\omega})$ for $q_{k,\omega} \leq \tilde{q}_k$. Appendix C.2.7 contains the Lagrangian equation and the first-order conditions. Proposition 6 determines the production quantities and the allocation of CEO time.

Proposition 6. *Suppose the firm produces in the headquarters and the establishment and incurs transport costs $\tau > 1$ to ship output from one location to the other. The firm allocates CEO time to equalize the marginal benefit of CEO time across headquarters and establishment. The marginal production costs typically differ. The firm chooses the production quantities either to equalize the marginal production costs adjusted by the transport costs across the headquarters and the establishment, or to produce local output. Formally, in optimum,*

$$\varphi_{0,\omega} = \varphi_{1,\omega}, \quad (26)$$

$$\tau\xi_{0,\omega} = \xi_{1,\omega} \qquad \textit{if } q_{0,\omega} = \tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega}), \quad (27)$$

$$\xi_{0,\omega} = \tau\xi_{1,\omega} \qquad \textit{if } q_{1,\omega} = \tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega}), \textit{ and} \quad (28)$$

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

$$\xi_{0,\omega} < \tau\xi_{1,\omega} \wedge \xi_{1,\omega} < \tau\xi_{0,\omega} \qquad \textit{if } q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0. \tag{29}$$

The marginal production costs are equal if establishments are symmetric, i.e. if $w_1 = w_0$*,* $\theta_{10} = \theta_{00}$*, and* $L_1 = L_0$ *(local output* $\tilde{q}_j$ *may differ).*

Proof. See Appendix C.2.8. □

The firm can flexibly allocate CEO time.As in the case of a global product market, the firm reallocates CEO time until its marginal benefit is equal at the headquarters and the establishment. The transport costs limit the flexibility of the allocation of output. The firm chooses between three options: it produces output locally, ships it from the other location or does both. If the marginal production costs at the establishment are equal to the marginal costs at the headquarters including the transport costs, the firm produces part of the establishment output locally and ships part of it from the headquarters (equation 27). The analogous result holds if the marginal costs at the headquarters and the marginal costs at the establishment including the transport costs are equal (equation 28). If the marginal costs at the headquarters are lower than the marginal costs including transport costs at the establishment and vice versa, the firm produces as much output locally as it would like to supply (equation 29). This is the case if establishments are symmetric. Finally, if the marginal costs including transport costs at the headquarters are lower than the marginal costs at the establishment (or vice versa), the firm produces total output in the headquarters (establishment), i.e. it is a single establishment firm. The cases described in equations (27) and (28) are evidently unstable because small changes in the marginal costs lead to a violation of the equations.

Corollary 3. *Suppose the firm incurs transport costs* $\tau > 1$ *to ship output from one location to the other. It can be optimal to produce at an establishment with the same number of below-CEO management layers as the headquarters* $L_0 = L_1$ *if the helping costs across space exceed those within a location,* $\theta_{10} > \theta_{00}$*, but the wages are equal or higher at the establishment than at the headquarters,* $w_1 \geq w_0$*.*

Proof. See Appendix C.2.8. □

As in the case of a global product market, production in the establishment is only efficient if some advantage counterbalances the higher helping costs across space θ_{10}. Avoiding the transport costs τ is an advantage that can make multi-establishment production with the same number of below-CEO layers optimal even if local wages at the establishment are higher than headquarter wages.

Comparative statics. The transport costs affect the comparative statics results. The results on local output $\tilde{q}_j$ are similar to those in Proposition 3 if the marginal costs adjusted for the transport costs are equal across locations and to those in Proposition 1 if they are not.

Proposition 7. *Suppose the firm produces in the headquarters and the establishment. Suppose further that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_0$, $\tilde{q}_1$.

b) If either $\xi_{1,\omega} = \tau\xi_{0,\omega}$ or $\xi_{0,\omega} = \tau\xi_{1,\omega}$, the marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\ell \leq L_0$, do not vary with local output $\tilde{q}_0$, $\tilde{q}_1$.

c) If $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$, the marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\ell \leq L_j$, increase with local output $\tilde{q}_0$, $\tilde{q}_1$ if the CEO spends a sufficient share of time on the location with the increase of output. Higher local output $\tilde{q}_j$ increases the number of production workers at the same location $n^0_{j,\omega}$ and decreases their number at the other location $n^0_{k,\omega}$, $k \neq j$.

Proof. See Appendix C.2.9. □

As in Propositions 1 and 3, higher local output leads to higher CEO knowledge. The impact of higher local output on the other endogenous variables depends on whether the marginal costs adjusted for transport costs are equal at the headquarters and the establishment. If the marginal costs adjusted for transport costs are equal, the firm reallocates output and the comparative statics results are similar to those for a multi-establishment firm that does not incur transport costs. In particular, the marginal benefit of CEO time and the below-CEO knowledge levels are constant, because the firm accommodates higher output by reallocating CEO time and output between the headquarters and the establishment.

If instead the marginal costs adjusted for transport costs differ between the headquarters and the establishment, the firm does not reallocate output. The comparative statics results are similar to those for a single-establishment firm (that, by definition, cannot reallocate output either). In particular, the firm hires more production workers at the location with higher output. It hires fewer production workers at the other location, because the higher CEO knowledge allows producing the same output with fewer workers. If the CEO spends a sufficiently high share of time on the location with the increasing output, the increase of the number of production workers there outweighs the decrease at the other location. The number of problems generated and communicated to the CEO increases. To satisfy the CEO time constraint, below-CEO knowledge levels increase. Correspondingly, the marginal benefit of CEO time rises.

Whether the marginal costs adjusted for the transport costs are equal between locations is also relevant for the effect of higher helping costs across space θ_{10} on headquarter organization.

Proposition 8. *Suppose the firm produces in the headquarters and the establishment with $\theta_{10} > \theta_{00}$. Suppose further that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$, the knowledge of the employees at all below-CEO layers $z^{\ell}_{1,\omega}$, $\ell \leq L_1$, and the marginal production costs $\xi_{1,\omega}$ at the establishment increase with the helping costs θ_{10}. The total number of production workers $\sum_{j=0}^{1} n^0_{j,\omega}$ and the number of production workers at the establishment $n^0_{1,\omega}$ decrease.

b) If either $\xi_{1,\omega} = \tau\xi_{0,\omega}$ or $\xi_{0,\omega} = \tau\xi_{1,\omega}$, the marginal benefit of CEO time $\varphi_{0,\omega}$, the knowledge of the employees at all below-CEO layers $z^{\ell}_{0,\omega}$, $\ell \leq L_0$, the number of production workers $n^0_{0,\omega}$ and the marginal production costs $\xi_{0,\omega}$ at the headquarters increase with the helping costs θ_{10}.

c) If $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$, the marginal benefit of CEO time $\varphi_{0,\omega}$, the knowledge of the employees at all below-CEO layers $z^{\ell}_{0,\omega}$, $\ell \leq L_0$, the number of production workers $n^0_{0,\omega}$ and the marginal production costs $\xi_{0,\omega}$ at the headquarters decrease with the helping costs θ_{10}.

These comparative statics results hold for $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $k \neq j$, if $L_0 \leq L_1$ or if $L_0 > L_1$ and wages w_1 are sufficiently low, and for $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $k \neq j$, if $L_1 \leq 1$ or if $L_1 = 2$ and the establishment's share of CEO time $s_{1,\omega}$ is sufficiently high.

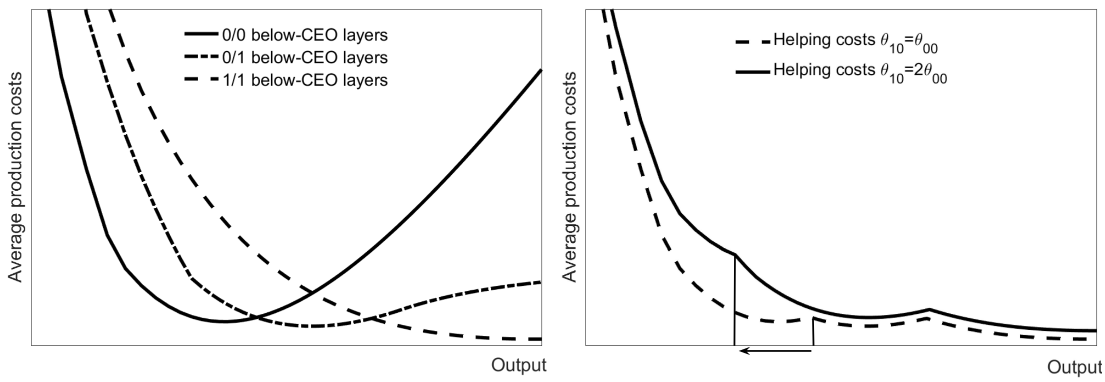
Proof. See Appendix C.2.10. □

As in the case of a global product market, higher helping costs θ_{10} make it more costly to use CEO knowledge. The firm increases CEO knowledge to compensate the higher costs with a higher benefit of using the CEO. As labor and CEO knowledge are complementary inputs, higher CEO knowledge allows the firm to decrease the total number of production workers. The firm decreases the number of production workers at the establishment, because the higher helping costs θ_{10} increase the costs of generating problems there. Due to the CEO time constraint, more problems have to be solved at below-CEO layers, so the knowledge of the employees in the establishment increases. In consequence, the marginal production costs rise.

The increase of the helping costs for the establishment θ_{10} affects the headquarters. 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

Figure 2: Illustration of the average cost functions, local product markets

(a) Symmetric locations (b) Helping costs θ_{10}



The figure plots the average cost function of a multi-establishment firm for $w_0 = w_1$, $\theta_{00} \leq \theta_{10}$, $\tilde{q}_0 = \tilde{q}_1$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$. (a): At each kink, the multi-establishment firm adds a layer at one establishment ($\theta_{10} = \theta_{00}$). (b): Higher helping costs θ_{10} decrease the total production quantity at which the firm reorganizes.

workers at the headquarters increases, as do their knowledge and the marginal production costs. If instead the marginal costs adjusted for the transport costs differ between the headquarters and the establishment, reallocating output is not efficient for the firm. The number of production workers at the headquarters decreases because local output is constant, but CEO knowledge increases. Correspondingly, the knowledge of the employees at the below-CEO layers at the headquarters decreases, as do the marginal production costs. The difference between the two cases is also reflected in the effect of higher helping costs θ_{10} on the marginal benefit of CEO time. Reallocating output increases the efficiency of the use of CEO time, so its marginal benefit increases. In contrast, producing strictly local output reduces the efficiency of the use of CEO time and thus its marginal benefit. In summary, whether reallocating output is efficient is decisive for how changes of establishment characteristics affect the organization of the headquarters.

Organizational structure. The firm chooses the organizational structure with the minimum production costs. Figure 2a illustrates the average production costs of different possible organizational structures if local output is equal and locations are symmetric, i.e., if wages and the helping costs are equal: $w_1 = w_0$, $\theta_{10} = \theta_{00}$. The average production costs are U-shaped, as in the case of a single-establishment firm. This reflects that reallocating output is only efficient under certain conditions. These conditions do not hold for symmetric locations if the number of below-CEO layers is equal, for example. As Proposition 7c) shows, the impact of higher output on firm organization is similar to its impact on a single-establishment firm if reallocating output is not efficient. This explains why the shape of the average cost function is similar to the single-establishment case. Importantly, though, the multi-establishment firm reorganizes gradually as it grows, as does a multi-establishment firm in a global product market. The firm does not add a layer at the headquarters and the establishment at the same size, but successively at one and another.

Figure 2b illustrates how the helping costs across space θ_{10} affect the number of managerial layers of the firm. Higher helping costs increase the production costs of the firm. The higher the helping costs are, the smaller is the quantity at which the firm adds a layer at the establishment, as a comparison of the solid and dashed lines show. 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 and thus marginal costs.

Importantly, the helping costs are not the only determinant of where the firm adds a layer. As Figure X in the Appendix shows, if wages at the headquarters are sufficiently lower than wages at the establishment, the firm may choose to add a layer at the headquarters instead of the establishment.

4.4 The optimal output

We return to the setting with many firms i that each produce a differentiated product outlined at the beginning of section 4.1. Firms compete monopolistically. Each firm is characterized by the taste draw for its product α_i. Agents maximize their utility (1) subject to their budget constraint. The total demand results from multiplying the individual demand by the number of agents per location:

$$q_j(\alpha_i) = \alpha_i R_j P_j^{\sigma-1} p_j(\alpha_i)^{-\sigma}$$

$R_j = N_j w_j$ denotes local income and P_j is the price index. We normalize P_0 to 1.

Each firm chooses the output levels to maximize profits:

$$\max_{\tilde{q}_0, \tilde{q}_1 \geq 0} \pi_i(\alpha_i) = \sum_{j=0}^{1} p_j(\tilde{q}_j(\alpha_i))\tilde{q}_j(\alpha_i) - C(\tilde{q}_0, \tilde{q}_1) \tag{30}$$

The optimal output is therefore equal to:

$$\tilde{q}_j(\alpha_i) = \alpha_i R_j P_j^{\sigma-1} \left(\frac{\sigma}{\sigma-1} \xi_{j,\omega} \left(\tilde{q}_0(\alpha_i), \tilde{q}_1(\alpha_i) \right) \right)^{-\sigma}, \tag{31}$$

where we make explicit that the marginal costs $\xi_{j,\omega}$ are a function of α_i through output. The optimal price is a constant mark-up over marginal costs:

$$p_j(\alpha_i) = \frac{\sigma}{\sigma-1} \xi_{j,\omega} \left(\tilde{q}_0(\alpha_i), \tilde{q}_1(\alpha_i) \right) \tag{32}$$

Proposition 9. *Suppose that the firm produces at the headquarters and the establishment. Suppose further that the local production quantities are equal to the amount of output that the firm supplies locally (i.e., $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$) and that production quantities are sufficiently large. Higher helping costs across space θ_{10} decrease the optimal output at the establishment $\tilde{q}_1(\alpha_i)$ and increases the optimal output at the headquarters $\tilde{q}_0(\alpha_i)$.*

Proof. See Appendix C.3. □

Firm geography affects the size of the firm's operations. Higher helping costs increase the marginal production costs at the establishment and decrease the marginal production costs at the headquarters if the firm produces local output locally. Correspondingly, higher helping costs decrease the optimal output at the establishment and increase the optimal output at the headquarters. If instead the firm produces part of local output at the other location, both the optimal output at the establishment and at the headquarters decrease, because the marginal production costs increase at both locations. As explained above, this case is unstable, however.

4.5 Summary

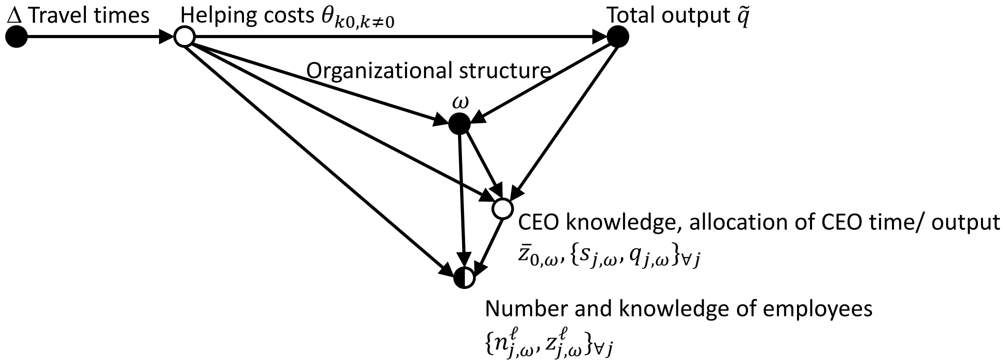
This section summarizes how the model explains the stylized facts in section 3.

Fact 1 documents that distance to the headquarters correlates negatively with the investment probability at a location and with establishment size. In the model, the helping costs θ_{10} proxy distance and other geographic frictions. Higher helping costs increase the marginal production costs of the establishment (Proposition 4, 8). They decrease the optimal output and thus the size of an establishment (Proposition 9) and the attractiveness of a location for an establishment (Corollaries 1-3).

Fact 2 shows that the number of managerial layers of a firm increases with the distance of its establishments, and that both the establishment's and the headquarters' organization respond. In the model, the helping costs θ_{10} not only affect the optimal choices at the establishment, but also at the headquarters due to the common CEO (Proposition 4, 8). The higher marginal costs increase the use of middle managers. Depending on local wages, middle managers are hired either at the headquarters or at the establishment (section 4.3.2).

Fact 3 documents that multi-establishment firms reorganize gradually as they grow. In the model, hiring middle managers at the establishment (or the headquarters) releases CEO time that is reallocated across locations. Therefore, efficiency increases throughout the firm, which reduces the need for middle managers at the headquarters (or establishment, see Proposition 5 and section 4.3.2).

Figure 3: Response of endogenous variables to change in the travel times



The graph illustrates the response of the endogenous variables to a change of the travel times according to the model in Section 4. The arrows denote causal relationships between the variables at the nodes. The node symbol • (○) denotes that a variable is (un)observable. ◐ denotes that a group of variables contains observable and unobservable variables.

5 Reorganization due to high-speed train routes

The key implication of the model is that multi-establishment firm organization is interdependent across the headquarters and the establishments. We exploit the opening of high-speed train routes in Germany (similar to Charnoz et al., 2015; Bernard et al., 2017) to study how an exogenous change of spatial frictions within firms affects the organization of headquarters and establishments. The new routes make it easier to travel between the headquarters and the establishments and thus reduce the costs to manage the establishments from the headquarters. In the terms of the model, they decrease the helping costs θ_{10}.

5.1 Model predictions

The model helps understand how changes of the helping costs θ_{10} affect multi-establishment firm organization. Figure 3 illustrates the model predictions using a directed graph. Solid circles denote variables that are observable and hollow circles denote variables that are unobservable in our data. The arrows denote causal links between variables. To keep the graph simple, we group variables by the steps of firm optimization and use semi-solid circles if only part of the group is observable.[22]

Lower travel times reduce the helping costs between an establishment k and the headquarters $\theta_{k0,k\neq0}$. Lower helping costs θ_{k0} increase the optimal total output $\tilde{q}$. The helping costs thus have direct and indirect effects on the organization of multi-establishment firms. Lower helping costs directly affect the organizational structure ω, because they reduce the optimal number of layers and affect the attrac-

[22]To recap, the number and knowledge of employees is determined at the establishment level, taking as given CEO knowledge, the allocation of CEO time and output that are determined at the firm level, in turn taking as given the organizational structure, which is determined in the last step of cost minimization. Optimal total output is determined in the profit maximization.

tiveness of multi-establishment production. 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 output $q_{j,\omega}$ depend directly on θ_{k0}, but also indirectly through $\tilde{q}$ and ω. The choice of 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.

The complexity of the relation between the helping costs θ_{k0} and the organizational outcomes has implications for the interpretation of the empirical estimates. The model predictions about the impact of a reduction of the helping costs at the establishment and headquarter level in Propositions 4 and 8 hold conditional on output and the organizational structure. These variables do not vary exogenously from the helping costs. We would need instruments for both variables. If we conditioned on total output or the organizational structure in an establishment-level regression, the estimation would entail a "bad control" problem (Angrist and Pischke, 2014, p. 214-7). Our empirical exercise therefore estimates the total—direct and indirect—effect of changes in the helping costs.

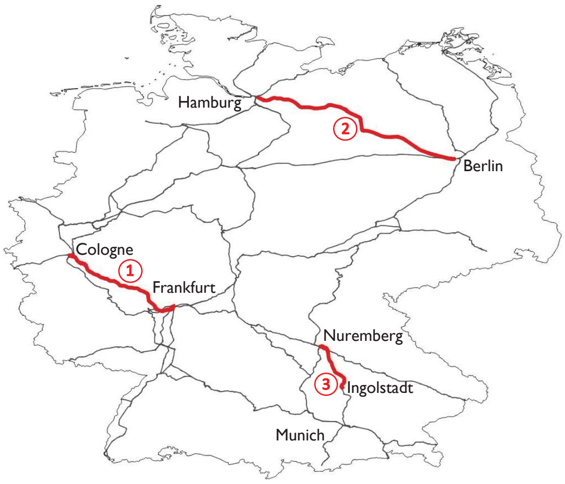
We focus on the model prediction that changes of the helping costs between an establishment and the headquarters not only affect the size and organization of this establishment, but also the size and organization of the headquarters (and possible other establishments) of the firm. In Appendix D.2, we show how it is possible to disentangle the direct and indirect effect of lower travel times on the organizational structure at the fir level by combining an estimate of the direct effect of lower travel times on output, their total effect on the organizational structure and an estimate of the effect of output on the number of managerial layers from Friedrich (2016).

5.2 Travel time data

We use information on the travel times between German cities from Deutsche Bahn AG, the state-owned German railway firm. We exploit that travel times changed substantially due to the opening of three high-speed train routes during our sample period.[23] Trains on all routes exclusively transport people. Figure 4 shows a map of the new high speed train routes and how they connect to the existing long distance rail network. Deutsche Bahn AG either constructed new rails (routes 1, 3) or substantially upgraded the existing railway network (route 2). Route 1 almost halved the travel time between Frankfurt and Cologne from 135 to 76 minutes. Service started in August 2002 (Eurailpress.de, 2002). Route 2 reduced the travel time between Hamburg and Berlin from 135 to 90 minutes from December 2004 (Eurailpress.de, 2004). Route 3 between Ingolstadt and Nuremberg opened in May 2006

[23] A forth route between Leipzig and Berlin opened in 2006. However, the travel time between Leipzig and Berlin decreased gradually according to the data, so we cannot use it in the estimation.

Figure 4: The new high speed train routes and the German long distance network



The map shows the German high speed rail network (black) and the new high speed train routes (bold red). Trains run at up to 300 km/h on the new routes, around 100 km/h faster than on the other routes. Data from Deutsche Bahn AG (http://data.deutschebahn.com/dataset/geo-strecke).

and reduced the travel time between the two cities from 66 to 30 minutes (Brux, 2006). Except for the Hamburg-Berlin connection, the high speed trains run at up to 300 km/h and thus around 100 km/h faster than on the other routes of the German long distance network.

As Figure 4 shows, the German railway network is very interconnected compared to other countries. Paris, for instance, is the center of the French railway network that has approximately a "star" network structure. The German railway network features several hubs. The travel time reductions therefore affect more locations than only those at the immediate ends. For example, route 1 between Cologne and Frankfurt decreased travel times from cities in the Ruhr area to many cities in East and South Germany such as Leipzig, Stuttgart, or Würzburg.

We use information on the mean and 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 total travel times as net travel times plus 30 minutes per change. Our data comprise 115 train stations that are connected to the long distance network in at least one of the three years. To make sure that temporary construction works do not affect the travel times, Deutsche Bahn AG computed the travel times for three different weekdays in March, June and November. Travel times may change over time for several reasons, like adjustments of the time table, 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% of passengers used one of the new high speed routes in 2008.

We merge the travel times and the data on multi-establishment firms using the information 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 that unobservable differences between firms connected and not connected to the network drive the results.

One may be worried that trains are not an attractive means of transportation for business travelers. However, the high-speed trains in particular are. According to information from Deutsche Bahn AG for the year 2017, the share of buisness travelers on the new routes was about double their average share.[24] This is not surprising given that the high speed routes make 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 than by train, for example—or even plane. In fact, regular plane service between Cologne Bonn Airport and Stuttgart Airport was discontinued in 2002,[25] and the service between Cologne Bonn Airport and Frankfurt Airport was discontinued in 2007.[26] The number of flights between Cologne Bonn Airport and Nuremberg Airport dropped substantially (Deutscher Bundestag, 2007).

5.3 Empirical specification

To understand the impact of faster travel times on directly affected establishments, we estimate:[27]

$$y_{ijt} = \beta_0 + \beta_1 D_{\theta_{jh}\downarrow,ijt} + \alpha_j + \alpha_{ct} + \epsilon_{ijt} \tag{33}$$

i refers to a multi-establishment firm, j to an establishment, h to the headquarters, c to the county where an establishment is located and t indexes time. α denotes fixed effects. The main variable of interest is $D_{\theta_{jh}\downarrow,ijt}$, an indicator variable for at least 30 minutes lower minimum travel times between the establishment and its headquarters.

To gauge indirect effects on the headquarters, we estimate:

$$y_{iht} = \beta_0 + \beta_1 D_{\exists j \text{ s.t. } \theta_{jh}\downarrow,iht} + \alpha_h + \alpha_{dt} + \epsilon_{iht} \tag{34}$$

where d denotes the headquarter county. To assess whether there are indirect effects on not directly affected establishments of affected firms, we estimate:

$$y_{ikt} = \beta_0 + \beta_1 D_{\Delta\theta_{kh}=0 \wedge \exists j \text{ s.t. } \theta_{jh}\downarrow,ikt} + \alpha_k + \alpha_{ct} + \alpha_{dt} + \epsilon_{ikt}, \quad k \neq j \tag{35}$$

k refers to a not directly affected establishment. The indicator $D_{\Delta\theta_{kh}=0 \wedge \exists j \text{ s.t. } \theta_{jh}\downarrow,ikt}$

[24]The statistics are computed based on the fraction of tickets sold with a corporate discount.

[25]It takes about two hours and 15 minutes to travel from Cologne to Stuttgart using the Cologne-Frankfurt high-speed route.

[26]The carrier Lufthansa cited the new high speed train route as main reason for lower demand (Eurailpress.de, 2007).

[27]This specification is similar to Charnoz et al. (2015).

is equal to one if the travel time between establishment k and the headquarters is constant, but the travel time between one of the other establishments of the firm and the headquarters decreases by at least 30 minutes. As outcome variables $y_{i.t}$, we use the number of non-managerial employees as measure for size and the number of managerial layers and the managerial shares for establishment organization.

The specifications mimic difference-in-differences estimation. The "treatment" is faster travel time between the directly affected establishment and the headquarters, or at least one establishment and the headquarters, respectively. Its baseline effect is captured by the establishment or headquarters fixed effects. The (headquarter) county × year fixed effects capture the "after" dummy. The indicator variables $D_{\theta_{jh}\downarrow,ijt}$, $D_{\exists j \text{ s.t. } \theta_{jh}\downarrow,iht}$ and $D_{\Delta\theta_{kh}=0\wedge\exists j \text{ s.t. } \theta_{jh}\downarrow,ikt}$ are equivalent to the interaction term of the "treatment" and "after" dummy variables. We implement the estimation using the `reghdfe` command by Correia (2014).

Difference in the travel times may also affect other model parameters, such as local wages because employees commute longer distances (Heuermann and Schmieder, 2018). Firms may also benefit from better suppliers (Bernard et al., 2017). The county-year and headquarter county-year fixed effects isolate the impact of lower spatial frictions on firm organization from other forces. Specifically, the regressions for directly affected establishments compare establishments with travel time reductions to their headquarters and establishments *in the same county and year* with constant travel times. Lower local wages or better suppliers benefit both types of establishments, so our estimation strategy differences out their effect. Similarly, the regressions for the headquarters compare headquarters with travel time reductions to at least one of their establishments to headquarters *in the same county and year* without. The specification for not directly affected establishments compares establishments that belong to firms with other treated establishments to establishments in the same county and year that belong to firms without treated establishments, additionally accounting for shocks at the headquarters location.[28] As being treated in this set-up presupposes that firms have at least two establishments in addition to their headquarter, we restrict the sample accordingly.

A possible concern with respect to our identification strategy is that firms are aware of the construction of high speed train routes prior to their opening, so they may strategically locate their establishments. Importantly though, while the location of the routes is predictable, their opening is not. For example, route 3 between Ingolstadt and Nuremberg was initially scheduled to open in 2003, later in 2004 and eventually only opened in mid-2006. Changes to establishment organization should

[28]The strictest specification would condition on county × headquarter county × year fixed effects, i.e. compare establishments of firms with travel time reductions for at least one establishment to establishments of firms without reductions in the same county with headquarters in the same headquarters county. However, there are too few such pairs in the sample to run these regressions.

only materialize after opening. We make sure that treated establishments exist at least one year before the route is opened. A few establishments and headquarters move from one county to another during the sample period. We use their original location for the main analyses and drop them from the sample in robustness checks.

We set the indicators $D_{\theta_{j0}\downarrow,ijt}$, etc. equal to one if the travel time between an establishment and the headquarters decreases by at least 30 minutes because the high-speed train routes decrease the travel times by at least 30 minutes.[29] As Appendix Table D.1 shows, virtually none of the non-high-speed-route connections exhibit lower travel times of 30 minutes or more. The threshold thus helps us to ensure that the reduction is indeed driven by the exogenous new routes instead of potentially endogenous demand-driven adjustments to the time table.

5.4 Regression results

Table 7 presents the regression results. Columns 1 to 4 contain results for all firms. Columns 5 to 8 restrict the sample to firms with at least two establishments in addition to their headquarters. The top panel contains the results for directly affected establishments, the middle panel those for headquarters and the bottom panel those for the indirectly affected establishments.

Faster travel times increase the size of the directly affected establishments. The number of non-managerial employees increases by about seven percent in both samples. Interestingly, this increase is not accompanied by an increase of the number of layers. Instead, the managerial shares tend to decrease, although not significantly. These results are consistent with higher establishment growth due to faster travel times. As the middle panel shows, faster travel times lead to organizational adjustments at the headquarters. While headquarter size stays constant, the number of managerial layers and the managerial share in the wage sum increase. This effect is particularly pronounced for headquarters of firms with at least two establishments. It is consistent with the idea that the faster travel times help firms manage their growing establishments from the headquarters. Quantitatively, the coefficient estimates are equivalent to an increase of the managerial share by four percent in the average firm (seven percent if the share is computed according to Blossfeld). The indirect impact of faster travel times goes beyond the headquarters. Establishments that do not directly benefit from faster travel times, but belong to firms that do, grow more slowly than establishments that neither benefit from faster travel times nor belong to firms that do. While their number of managerial layers and the layer-based managerial share does not increase, the Blossfeld-defined managerial share

[29] One may be worried 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 high speed routes, 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 below).

Table 7: Regression results, 2000-2010 panel

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.074***	0.006	−0.029	−0.145	0.067***	0.000	−0.250	−0.147
	(0.011)	(0.009)	(0.246)	(0.143)	(0.012)	(0.009)	(0.262)	(0.154)
R-squared	0.890	0.859	0.832	0.868	0.891	0.864	0.833	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	94,354	94,354	94,354	94,354	83,894	83,894	83,894	83,894
# est.	13,544	13,544	13,544	13,544	12,244	12,244	12,244	12,244
Headquarters								
Firm treated	−0.020	0.018	0.537*	0.063	−0.013	0.042*	0.996**	0.631*
	(0.014)	(0.015)	(0.257)	(0.192)	(0.019)	(0.019)	(0.320)	(0.251)
R-squared	0.945	0.882	0.926	0.892	0.950	0.889	0.931	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	22,884	22,884	22,884	22,884	12,264	12,264	12,264	12,264
# HQ	2,875	2,875	2,875	2,875	1,587	1587	1,587	1587
Not directly affected establishment								
Firm treated					−0.030**	0.004	0.221	0.412**
					(0.011)	(0.008)	(0.235)	(0.140)
R-squared					0.898	0.867	0.834	0.873
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

increases. This is consistent with a constant number of upper level managers in the establishments despite their smaller size. Overall, the results strongly support the model implication that multi-establishment firm organization is interdependent across headquarters and establishments.

Robustness. Appendix D documents the robustness of the results. Table D.2 replicates the regressions after dropping establishments or headquarters that move from one county to another. The results for the directly affected establishments are virtually unchanged; the results for the headquarters become even a little stronger. The regression results concerning the not directly affected establishments do not reflect the size decrease any more; this may stem from the considerably smaller sample size. Table D.3 replaces the county/headquarter county-year fixed effects with bundesland-year fixed effects.[30] The regressions contain both establishment and county/headquarter county-year fixed effects. One may be concerned that estimating so many fixed effects off a limited-size sample may lead to an "overfitting" of the data. The bundesland-year fixed effects reduces the number of geographic fixed effects substantially from up to 1,500 to less than 180. The estimated effects tend to be larger and slightly more significant. We conclude that the results in Table 7 are unlikely to be an artifact of overfitting the data. Table D.4 documents that the results are robust to defining the treatment dummies based on the change of mean instead of minimum travel times between locations. While the main specification employs robust standard errors, Table D.5 shows that the main effects are robust to clustering standard errors by establishment (or headquarters, respectively) and (headquarter) county. Only for the headquarters, effects are now significant only in the the sample of firms with at least two establishments.

Specifically for the directly affected establishments, Table D.6 applies the within-firm-across-establishments identification strategy of Giroud (2013). Results are robust. An important caveat is that they are identified from only 150 distinct firms, because there are only few firms with treated and untreated establishment. Finally, we restrict the sample to stations that are connected to the long distance network in all years to ensure that the high speed routes drive the travel time changes instead of connecting stations to the network. Table D.7 shows that the results are robust.

6 Conclusion

This paper studies how geographic frictions affect multi-establishment firm organization. We show that geographic frictions increase the use of middle managers, and that geographic frictions between the headquarters and one establishment not only affect the organization of the establishment, but also of the headquarters and other establishments of a firm. An important implication of our study is that local conditions propagate across space through firm organization.

Advances in information and communication technologies are likely to decrease geographic frictions. This affects the organization of existing multi-establishment firms and, at the same time, renders multi-establishment production attractive for an increasing number of firms. Understanding the role of geographic frictions for firm performance thus remains an exciting area for future research.

[30] We do not apply this variation to the indirectly affected establishments, because headquarters and establishments are often in the same bundesland.

References

Angrist, J. D. and J.-S. Pischke (2014). *Mastering 'metrics: The path from cause to effect*. Princeton University Press.

Antoni, M., A. Ganzer, and P. vom Berge (2016). Sample of integrated labour market biographies (SIAB) 1975-2014. FDZ-Datenreport 04/2016.

Antoni, M., K. Koller, M.-C. Laible, and F. Zimmermann (2018). Orbis-ADIAB: From record linkage key to re-search dataset. Combining commercial company data with administrative employer-employee data. FDZ-Methodenreport, forthcoming.

Antoni, M. and R. Schnell (2017). The past, present and future of the German Record Linkage Center (GRLC). *Jahrbücher für Nationalökonomie und Statistik forthcoming*.

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

Battiston, D., J. Blanes i Vidal, and T. Kirchmaier (2017). Face-to-Face Communication in Organisations. *mimeo, London School of Economics*.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2017). Production networks, geography and firm performance. *Journal of Political Economy forthcoming*.

Bertrand, M. (2009). Ceos. *Annual Review of Economics 1*(1), 121–150.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2016). What drives differences in management? *mimeo, Stanford University*.

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

Blossfeld, H.-P. (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, G. (2006). NBS Nürnberg - Ingolstadt in Betrieb. *Der Eisenbahningenieur* (June), 142–143.

Bundesagentur für Arbeit (2016). Arbeitsmarkt in Zahlen: Beschäftigungsstatistik.

Caliendo, L., G. Mion, L. D. Opromolla, and E. Rossi-Hansberg (2015a). Productivity and organization in Portuguese firms. *NBER Working Paper No. 21811*.

Caliendo, L., F. Monte, and E. Rossi-Hansberg (2015b). The Anatomy of French Production Hierarchies. *The Journal of Political Economy 123*(4), 809–852.

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

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

Chandler, Jr, A. D. (2002). *The visible hand.* Cambridge, MA/London: Harvard University Press.

Charnoz, P., C. Lelarge, and C. Trevien (2015). Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail. *Economic Journal forthcoming.*

Chen, C. (2017). Management Quality and Firm Hierarchy in Industry Equilibrium. *American Economic Journal: Microeconomics 9*(4), 203–44.

Chen, C. and W. Suen (2017). The Comparative Statics of Optimal Hierarchies. *mimeo, University of Hong Kong*.

Christen, P. (2012). *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection.* Berlin: Springer.

Correia, S. (2014). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

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

Deutscher Bundestag (2007). Flughafenanbindungen von Fernverkehrshaltepunkten der Bahn. BT-Drucksache 16/7104.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009). Revisiting the German wage structure. *The Quarterly Journal of Economics 124*(2), 843–881.

Eberle, J., P. Jacobebbinghaus, J. Ludsteck, and J. 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.

Eurailpress.de (2002, July). NBS Köln-Rhein/Main wird heute eröffnet. http://www.eurailpress.de/news/alle-nachrichten/single-view/news/nbs-koeln-rheinmain-wird-heute-eroeffnet.html, accessed on 24 January 2018.

Eurailpress.de (2004, December). Hamburg - Berlin: Hochgeschwindigkeitsstrecke feierlich eröffnet. http://www.eurailpress.de/news/alle-nachrichten/single-view/news/hamburg-berlin-hochgeschwindigkeitsstrecke-feierlich-eroeffnet-1.html, accessed on 24 January 2018.

Eurailpress.de (2007, September). Schnelle Bahnverbindung: Lufthansa streicht Flüge Köln Frankfurt. http://www.eurailpress.de/news/alle-nachrichten/single-view/news/schnelle-bahnverbindung-lufthansa-streicht-fluege-koeln-frankfurt-1.html, accessed on 24 January 2018.

Fitzenberger, B., A. Osikominu, and R. Völter (2005). Imputation rules to improve the education variable in the IAB employment subsample. *Schmollers Jahrbuch 126*(3), 405–436.

Friedrich, B. (2016). Trade shocks, firm hierarchies and wage inequality. *mimeo, Northwestern University*.

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

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

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

Giroud, X. and H. M. Mueller (2017). Firms' Internal Networks and Local Economic Shocks. Technical report, National Bureau of Economic Research.

Guadalupe, M. and J. 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, A. (2018). The organization of knowledge in multinational firms. *Journal of the European Economic Association forthcoming.*

Hall, A. and M. 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.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004). Export versus FDI with heterogeneous firms. *The American Economic Review 94*(1), 300–316.

Hethey, T. and J. Schmieder (2010). Using Worker Flows in the Analysis of Establishment Turnover Evidence from German Administrative Data. FDZ-Methodenreport 06/2010.

Heuermann, D. F. and J. F. Schmieder (2018). The effect of infrastructure on worker mobility: evidence from high-speed rail expansion in Germany. *Journal of Economic Geography*.

Irarrazabal, A., A. Moxnes, and L. D. Opromolla (2013). The margins of multinational production and the role of intrafirm trade. *Journal of Political Economy 121*(1), 74–126.

Jacobebbinghaus, P. and S. Seth (2007). The German integrated employment biographies sample IEBS. *Schmollers Jahrbuch 127*(2), 335–342.

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

Mariscal, A. (2018). Firm Organization and Information Technology: Micro and Macro Implications.

Menor, L. and K. Mark (2001). Blinds To Go: Invading the Sunshine State. Ivey Publishing.

Nueno, P. and P. Ghemawat (2002). Revitalizing Philips.

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

Schild, C.-J. (2016). Linking "Orbis Company" Data with Establishment Data from the German Federal Employment Agency. German RLC Working Paper No. wp-grlc-2016-02, http://soz-159.uni-duisburg.de/wp-content/uploads/2017/07/wp-grlc-2016-02.pdf.

Schmucker, A., S. Seth, J. Ludsteck, J. Eberle, and A. Ganzer (2016). The Establishment History Panel 1975-2014. FDZ-Datenreport 03/2016.

Seetharam, I. (2018). The Indirect Effects of Hurricanes: Evidence from Firm Internal Networks.

Sforza, A. (2017). Shocks and the organization of the firm: Who pays the bill?

Appendix

A Data

A.1 Data sources and record linkage procedure

A.1.1 Social security records

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 cover 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. Jacobebbinghaus and Seth (2007) and Antoni et al. (2016) provide a detailed description of the structure of the data.[31]

Information on education is not reported for all periods for every individual, but can be inferred from other observations on the same individual. We follow the imputation procedure in Fitzenberger et al. (2005) and impute missing values for 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.[32] 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 file entry and exit. The extension file entry and exit uses information on worker flows to identify establishment openings and closings. Establishment identifiers may change when a firm restructures. The extension file helps mitigate bias related to restructurings. Hethey and Schmieder (2010) provide details on the file.

A.1.2 Orbis

We use a linkage table between the social security records and the firm-level database Orbis of the commercial data provider Bureau van Dijk (BvD). BvD compiles its firm-level 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.

It is important to note that BvDs financial information on firms in Germany is most reliable since 2006, as there have been some changes in the financial reporting system in Germany in that year. In earlier years, a higher share of financial information is missing.

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

[32] 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.

A.1.3 Record linkage procedure

The record linkage between Orbis and the social security data was performed independently of our project by the German Record Linkage Center (GRLC, see Antoni and Schnell (2017) or www.record-linkage.de for more details on the GRLC). 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 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, legal form and address 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 only 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 of the 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 multi-establishment 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 under consideration. Antoni et al. (2018) provide details on the process.

A.2 Sector and occupation classification

The information on the establishment sector changes over time. The sector information uses the respective latest sector classification of the German Statistical Office that updated the classification in 1993, 2003 and 2008. We follow Eberle et al. (2011) and transfer the sector classification after 2003 into the classification as of 1993.

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. (2015b) 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 (see section A.2). 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 (2016), we generally assign the minimum level of layers to these occupations. Table A.1 displays our assignment of occupations to layers.

Managerial occupations according to Blossfeld (1983, 1987). The assignment 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	63124, 71104, 73294, 84394, 94494	Manager, executive, director, board member
2	721, 722, 724, 752, 753, 761, 763, 843	All sub-groups of type 2 in occupation groups: 434, 524, 815; of type 3 in occupation groups: 411, 431, 434, 524, 922; of type 4 in occupation groups: 115, 411, 412, 431, 432, 433, 434, 511, 513, 516, 524, 532, 621, 625, 632, 633, 634, 712, 713, 715, 722, 723, 731, 732, 815, 824, 921, 922, 933; plus: 11494, 21194, 23294, 27194, 27294, 27394, 29194, 29294, 31174, 31194, 41203, 41303, 41383, 41304, 41384, 41394, 41403, 41404, 41484, 41494, 42124, 42144, 42314, 42324, 42394, 43152, 43323, 43343, 43353, 43383, 51133, 51233, 51533, 51543, 51594, 53184, 53394, 61194, 61294, 61394, 63114, 63194, 63313, 71224, 71333, 71433, 72144, 72194, 72243, 73394, 81214, 81234, 81404, 81414, 81424, 81434, 81444, 81454, 81464, 81474, 81484, 81804, 81814, 81884, 82594, 83193, 83194, 84194, 84294, 84304, 84494, 91344, 91354, 92113, 92304, 92394, 92424, 92434, 93303, 93313, 93323, 93343, 93383, 94214, 94493, 94404, 94414, 94484, 94534, 94794	Manager in business organization and strategy, finanical analyst, software developer, qualified IT-specialist, lawyers
1	31, 32, 601, 602, 603, 605, 606, 607, 611, 612, 621, 622, 623, 625, 626, 627, 628, 629, 762, 811, 813, 841, 842, 844, 862, 863, 871, 872, 873, 874, 875, 881,	All sub-groups of type 2 in occupation groups: 271, 273, 311, 312, 412, 414, 421, 613, 634, 811, 812, 817, 818, 821, 833, 844, 931, 932, 944, 946, 947; of type 3 in occupation groups: 233, 271, 312, 341, 421, 422, 423, 432, 523, 531, 532, 533, 541, 611, 612, 613, 625, 634, 721, 723, 733, 811, 812, 816, 817, 818, 821, 822, 833, 842, 845, 912, 913, 923, 924, 931, 941, 942, 945, 946, 947; of type 4 in occupation groups: 117, 221, 222, 223, 231, 233, 234, 241, 242, 243, 244, 245, 251, 252, 261, 262, 263, 312, 321, 322, 341, 342, 343, 422, 512, 523, 714, 813, 816, 817, 821, 822, 833, 845, 911, 912, 914, 931, 932, 935, 936, 941, 943, 946;	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
	882, 883, 604, 624, 633, 687, 812, 822, 831, 851, 852, 853, 855, 891, 892, 893, 922	plus: 1104, 11132, 11103, 11113, 11123, 11133, 11104, 11114, 11124, 11184, 11233, 11214, 11423, 11424, 11603, 11604, 11713, 11723, 12103, 12113, 12123, 12104,12144, 21113, 21114, 21124, 21213, 21223, 21233, 21313, 21323, 21363, 21413, 21423, 22103, 22183, 22222, 22203, 22303, 22333, 22343, 23113, 23123, 23222, 23223, 23224, 23322, 23413, 23423, 24133, 24203, 24233, 24303, 24413, 24423, 24513, 24523, 24533, 25103, 25133, 25183, 25213, 25223, 25233, 25243, 25253, 26113, 26123, 26223, 26243, 26253, 26263, 26303, 26313, 26323, 26333, 26383, 27104, 27184, 27212, 27223, 27283, 27224, 27284, 27313, 27304, 27314, 28103, 28113, 28123, 28133, 28143, 28104, 28114, 28213, 28223, 28214, 28224, 28313, 28343, 28314, 29103, 29113, 29123, 29133, 29143, 29104, 29114, 29134, 29203, 29213, 29223, 29233, 29243, 29253, 29263, 29273, 29283, 29204, 29284, 31103, 31133, 31143, 31153, 31163, 31173, 31104, 31114, 31124, 31134, 31144, 31154, 31164, 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, 41322, 41313, 41323, 41314, 41324, 41413, 41423, 41433, 41483, 41414, 41424, 41434, 42114, 42134, 42202, 42334, 43102, 43112, 43122, 43313, 43333, 43363, 51182, 51113, 51123, 51183, 51223, 51243, 51503, 51513, 51523, 51583, 51593, 51504, 51534, 51623, 51663, 53152, 53124, 53134, 53222, 53232, 53312, 53322, 53332, 53314, 61132, 61124, 61204, 61214, 61284, 61314, 62183, 63122, 63132, 63123, 63212, 63213, 71403, 71423, 71522, 71523, 72124, 72134, 72184, 72213, 72223, 72233, 73162, 73163, 73183, 73241, 73202, 73212, 73232, 73242, 73282, 73203, 73213, 73233, 73243, 73253, 73283, 73314, 73324, 73334, 81224, 81294, 81302,81332, 81352, 81382, 81313, 81323, 81333, 81353, 81383, 81393, 81494, 81894, 82212, 82232, 82332, 82343, 82522, 82503, 82523, 82504, 82514, 82524, 82534, 83112, 83132, 83123, 83133, 83124, 83134, 83154, 83223, 84114, 84124, 84134, 84144, 84184, 84214, 84224, 84413, 84404, 84414, 84424, 84434, 84444, 84454, 84484, 91314, 91324, 91334, 91384, 92133, 92384, 92414, 92494, 93213, 93223, 93233, 93333, 93413, 93433, 93513, 93523, 93603, 93613, 93623, 93633, 93643, 93653, 93683, 94224, 94303, 94313, 94323, 94403, 94413, 94483, 94522, 94532, 94582, 94514, 94704, 94714, 94724	

Continued on next page

Table A.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
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 Descriptive evidence on occupations by layer

Table A.2: Log wages and education by layer

Layer	N	Mean	p25	p50	p75
		Log wages			
0	4,313,387	4.441	4.181	4.453	4.715
1	1,396,129	4.765	4.533	4.816	5.102
2	594,340	4.965	4.841	5.089	5.173
3	52,216	5.059	5.018	5.173	5.173
		Education			
0	4,313,387	2.010	2	2	2
1	1,396,129	2.744	2	2	4
2	594,340	3.243	2	4	4
3	52,216	3.228	2	4	4

2012 cross-section, only firms with at least 10 employees. *Log wages:* log daily wages. The median and 75th percentile are equal for employees at layer 3, because wages exceed the social security limit. *Education:* 1 - Primary school/ lower secondary school/ intermediate school leaving certificate, no vocational qualification; 2 - Primary school/ lower secondary school/ intermediate school leaving certificate, with vocational qualification; 3 - Upper secondary school leaving certificate (Abitur), with or without vocational qualification; 4 - Degree from university/ university of applied sciences. Share of employees in category 4 by layer: 0 - 6.2%, 1 - 34.0%, 2 - 60.5%, 3 - 60.7%.

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 contains information on the occupation of employees. We relate the tasks of employees to the layer assigned their occupation by estimating, via OLS:

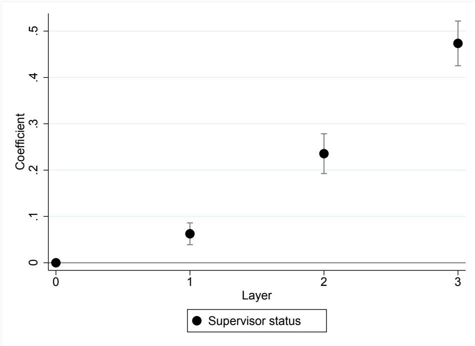
$$y_i = \beta \mathbf{D}_{\text{layer},i} + \gamma \mathbf{X}_i + \delta \mathbf{Z}_i + u_i \tag{A.1}$$

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.

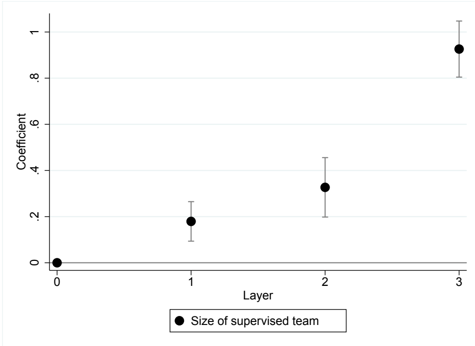
Figure A.1 plots the coefficients and confidence bands by layer. Employees at higher layers are significantly more likely to be supervisors. The predicted probability that an employee at layer 3 is a supervisor at the mean is 84%. Employees at higher layers also supervise larger teams. They are more likely to independently organize their own work. Their duties comprise organizing work for others, making decisions and solving problems. The job of employees at higher layers also require more specific skills. Overall, this descriptive evidence is consistent with the assumption that the assignment of occupations to layers reflects differences between the managerial tasks and duties of employees in firms. Table A.3 presents the estimated coefficients.

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

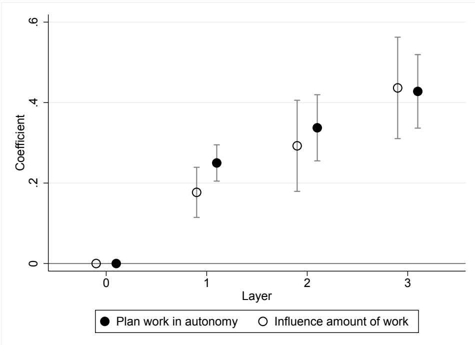
(a) Supervisor status



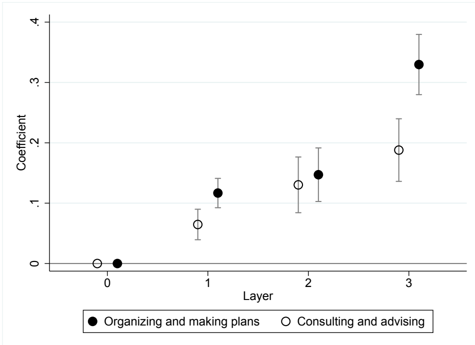
(b) Team size



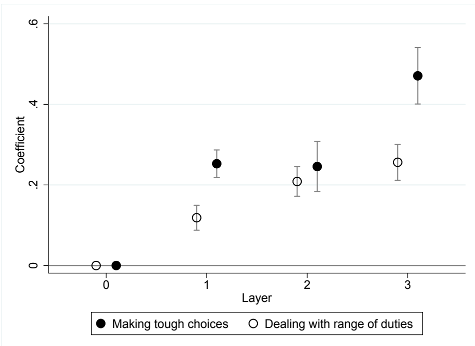
(c) Freely organizing own work



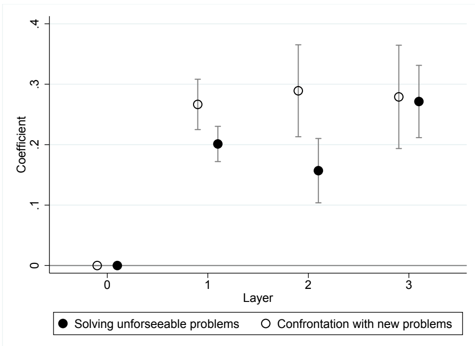
(d) Organizing others' work



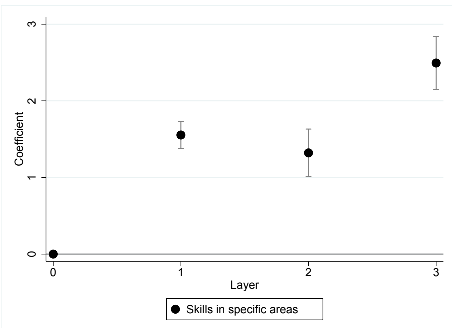
(e) Decision making



(f) Facing and solving problems



(g) Skills



The figure plots the estimated coefficients of the layer dummies in equation (A.1) for different survey questions. See Table A.3 for the survey questions.

Table A.3: Regression results: tasks by layer, 2006 BiBB/BAuA survey

	(a)	(b)	(c1)	(c2)	(d1)	(d2)	(e1)	(e2)	(f1)	(f2)	(g)
Layer 1	0.063***	0.179***	0.250***	0.177***	0.117***	0.199***	0.253***	0.119***	0.201***	0.267***	1.554***
	(0.012)	(0.044)	(0.023)	(0.031)	(0.012)	(0.012)	(0.017)	(0.015)	(0.015)	(0.021)	(0.090)
Layer 2	0.236***	0.327***	0.337***	0.293***	0.147***	0.252***	0.246***	0.209***	0.157***	0.289***	1.321***
	(0.022)	(0.066)	(0.042)	(0.057)	(0.023)	(0.022)	(0.032)	(0.018)	(0.027)	(0.038)	(0.159)
Layer 3	0.474***	0.926***	0.428***	0.436***	0.330***	0.327***	0.471***	0.256***	0.271***	0.279***	2.494***
	(0.025)	(0.062)	(0.047)	(0.063)	(0.025)	(0.025)	(0.036)	(0.022)	(0.031)	(0.043)	(0.177)
Age	0.000	0.003	−0.002	0.005***	−0.003***	−0.003***	−0.002***	−0.003***	−0.007***	−0.006***	−0.026***
	(0.000)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Tenure	0.037***	0.088***	0.087***	0.021	0.018***	0.027***	0.054***	0.027**	0.001	0.050***	0.151***
	(0.005)	(.020)	(0.010)	(0.014)	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.009)	(0.043)
Gender	−0.114***	−0.183***	0.028	−0.187***	−0.021*	0.100***	−0.169***	0.027	−0.120***	−0.082***	−1.282***
	(0.009)	(0.035)	(0.017)	(0.023)	(0.009)	(0.009)	(0.013)	(0.014)	(0.011)	(0.015)	(0.072)
Constant	0.009	0.664	2.434***	2.197***	0.293***	0.507***	1.616***	1.624***	1.830***	2.223***	20.451***
	(0.036)	(0.210)	(0.121)	(0.163)	(0.037)	(0.036)	(0.052)	(0.057)	(0.044)	(0.062)	(0.296)
# observations	12,514	4,400	11,958	11,926	12,514	12,514	12,510	12,509	12,511	12,510	10,282

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Regression results of equation A.1. *Dependent variables* defined by questions from BiBB survey: (a) Supervisor status (Y/N); (b) How many people do you supervise?; (c1) You are allowed to plan and schedule your work by yourself; (c2) You are able to influence the amount of work you have to do; (d1) How frequently are you organizing, making plans, working out operations?; (d2) How frequently are you consulting, advising?; (e1) Making tough choices on own responsibility; (e2) Dealing with a range of duties and responsibilities; (f1) Having to react to and solving unforeseeable problems; (f2) You are confronted with new problems that remain to be understood/familiarized with; (g) Skills in specific subject areas. *Independent variables: Layer X:* dummy variable for layer X; *Age:* age of survey participant in years; *Tenure:* tenure of survey participant in decades; *Gender:* gender of survey participant, 1=female. Education, firm size and sector fixed effects included.

B Facts

B.1 Distance to headquarters decreases location probability

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

Dependent variable	Location probability			Log # est. employees		
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	-0.303^{***}	-0.302^{***}	-0.363^{***}	-0.106^{***}	-0.105^{***}	-0.141^{***}
	(0.020)	(0.022)	(0.018)	(0.020)	(0.020)	(0.018)
Log market potential	0.649^{***}	0.672^{***}		0.508^{***}	0.510^{***}	
	(0.025)	(0.028)		(0.057)	(0.048)	
Relative wages	-0.794^{***}	-0.568^{***}		-0.375^{*}	-0.357^{*}	
	(0.062)	(0.064)		(0.160)	(0.148)	
Relative land prices		-0.019^{**}			0.010	
		(0.007)			(0.007)	
# of observations	xxx	xxx	xxx	181,336	91,885	181,336
# of firms	11,602	11,303	11,602	4,045	3,200	4,045
Year fixed effects	Y	Y	Y	N	N	N
Legal form fixed effects	Y	Y	Y	N	N	N
HQ sector fixed effects	Y	Y	Y	N	N	N
HQ county fixed effects	Y	Y	Y	N	N	N
County fixed effects	N	N	Y	N	N	Y
Firm-year fixed effects	N	N	N	Y	Y	Y

The table presents the coefficient estimates of a probit model (constant included; 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 3.

B.2 Distance to headquarters increases number of layers

Table B.2: Regression results, ME firm organization, 2012 cross-section, OLS

Dependent variable	# mgmt. layers				Mg. share ∈ [0, 1]			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.029***		0.033***		0.010***		0.002*	
	(0.008)		(0.006)		(0.002)		(0.001)	
Log area		0.042***		0.050***		0.014***		0.006***
		(0.011)		(0.009)		(0.002)		(0.001)
Log sales	0.262***	0.217***						
	(0.007)	(0.013)						
Log # non-mg. employees			0.287***	0.242***				
			(0.008)	(0.013)				
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
# firms	5,082	1,532	9,264	2,676	9,264	2,676	9,264	2,676

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 4. The number of observations is lower than the number of multi-establishment firms due to missing values for the legal form.

Table B.3: Regression results, establishment organization, 2012 cross-section, OLS

	Establishment			Headquarters		
Dependent variable	# layers	Mg. share (%)		# layers	Mg. share (%)	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.020***	0.410+	0.554*			
	(0.006)	(0.226)	(0.221)			
Maximum log distance to HQ				0.072***	2.160***	0.534***
				(0.007)	(0.189)	(0.128)
Log # non-mg. employees	0.301***			0.329***		
	(0.007)			(0.007)		
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
# est./HQ	35,061	35,061	35,061	9,802	9,802	9,802

The table presents the estimated coefficients. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable and independent variables:* see Table 5.

Table B.4: Regression results, ME firm organization, 2012 cross-section, non-linear size

Dependent variable	# mgmt. layers, Poisson			
	(1)	(2)	(3)	(4)
Maximum log distance to HQ	0.014***		0.019***	
	(0.004)		(0.004)	
Log area		0.017***		0.026***
		(0.004)		(0.004)
Log sales	0.508***	0.525***		
	(0.032)	(0.045)		
Log sales, squared	−0.018***	−0.019***		
	(0.001)	(0.002)		
Log # non-mg. employees			0.179***	0.238***
			(0.015)	(0.026)
Log # non-mg. employees, squared			−0.005***	−0.014***
			(0.002)	(0.002)
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y
# firms	5,111	1,661	9,275	2,768

The table presents the estimated coefficients. Constant included. Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with at least two establishments (plus headquarters). *Dependent variable:* number of managerial layers. *Independent variables:* see Table 4. The number of observations is lower than the number of multi-establishment firms due to missing values for the legal form.

Table B.5: Regression results, establishment organization, 2012 cross-section, non-linear size

Dependent variable	# layers	
	Est.	HQ
	(1)	(2)
Log distance to HQ	0.021**	
	(0.007)	
Maximum log distance to HQ		0.045***
		(0.004)
Log # non-mg. employees	0.456***	0.275***
	(0.023)	(0.014)
Log # non-mg. employees, squared	−0.027***	−0.014***
	(0.003)	(0.002)
Sector FE	Y	Y
County FE	Y	Y
# est./HQ	35,079	9,812

The table presents the estimated coefficients. Constant included. Standard errors in parentheses (clustered by firm in column 1, robust in column 2). ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* number of managerial layers. *Independent variables:* see Table 5.

Table B.6: Regression results, ME firm organization, 2012 cross-section, number of establishments

Dependent variable	# mgmt. layers, Poisson				Mg. share ∈ [0, 1], GLM			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.018***		0.019***		0.051***		0.029*	
	(0.004)		(0.004)		(0.008)		(0.012)	
Log area		0.023***		0.027***		0.075***		0.074***
		(0.005)		(0.004)		(0.011)		(0.014)
Log sales	0.125***	0.095***						
	(0.004)	(0.005)						
Log # non-mg. employees			0.141***	0.111***				
			(0.004)	(0.006)				
# establishments	−0.000**	−0.000+	−0.000	−0.000	−0.001	−0.001	0.001	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
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
# firms	5,111	1,661	9,275	2,768	9,275	2,768	9,275	2,768

The table presents the estimated coefficients. 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 at least two establishments (plus headquarters). *Dependent variable:* 1-4 number of managerial layers, 5-6 managerial share in wage sum, defined by layer, 7-8 managerial share in wage sum, defined by Blossfeld. *Independent variables: # establishments:* number of establishments (excluding HQ), others: see Table 4.

Table B.7: Regression results, ME firm organization, 2012 cross-section, distance quartiles

Dependent variable	# mgmt. layers, Poisson		Mg. share ∈ [0, 1], GLM	
			Layers	Blossfeld
	(1)	(2)	(3)	(4)
Quartile 2	0.021	0.014	−0.047	−0.055
	(0.019)	(0.015)	(0.033)	(0.048)
Quartile 3	0.079***	0.068***	0.126***	0.081+
	(0.018)	(0.015)	(0.034)	(0.048)
Quartile 4	0.083***	0.093***	0.271***	0.188***
	(0.018)	(0.015)	(0.035)	(0.049)
Log sales	0.124***			
	(0.004)			
Log # non-mg. employees		0.138***		
		(0.004)		
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y
# firms	5,111	9,275	9,275	9,275

The table presents the coefficients. 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, defined by layer, (4) managerial share in wage sum, defined by Blossfeld. *Independent variables: Quartile 2-4*: dummies for quartiles of the log of maximum distance between establishment and headquarters in km; others see Table 4.

Table B.8: Regression results, establishment organization, 2012 cross-section, distance quartiles

	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 quartile 2	0.076***	0.093*	0.107			
	(0.020)	(0.038)	(0.091)			
Log distance quartile 3	0.097***	0.142**	0.310*			
	(0.023)	(0.048)	(0.138)			
Log distance quartile 4	0.087***	0.114*	0.413**			
	(0.025)	(0.049)	(0.128)			
Max. log distance quartile 2				0.087***	0.117**	0.084
				(0.018)	(0.038)	(0.058)
Max. log distance quartile 3				0.137***	0.285***	0.180**
				(0.018)	(0.038)	(0.057)
Max. log distance quartile 4				0.190***	0.503***	0.372***
				(0.018)	(0.039)	(0.056)
Log # non-mg. employees	0.309***			0.182***		
	(0.012)			(0.004)		
Model	Poisson	GLM	GLM	Poisson	GLM	GLM
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
# est./HQ	35,079	35,079	35,079	9,812	9,812	9,812

The table presents the coefficients. 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:* (1),(4) number of managerial layers, (2),(5) managerial share in wage sum, defined by layer, (3),(6) managerial share in wage sum, defined by Blossfeld. *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 subordinate establishment and headquarters in km; *Log # of non-mg. employees:* log number of employees at lowest layer in establishment/HQ.

Table B.9: Regression results, ME firm organization, firm level, 2012 cross-section

	# mgmt. layers, Poisson				Managerial share Layers	
	(1)	(2)	(3)	(4)	(5)	(6)
Maximum log distance to HQ	0.011***		0.021***		0.639**	
	(0.004)		(0.003)		(0.215)	
Log area		0.010***		0.012***		0.151
		(0.003)		(0.002)		(0.162)
Log sales	0.115***	0.090***				
	(0.004)	(0.005)				
Log # non-mg. employees			0.115***	0.090***		
			(0.003)	(0.005)		
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
# firms	5,039	1,984	9,287	3,320	9,275	3,320

2012 cross-section. Even columns include only ME firms with at least two subordinate establishments. Robust standard errors in parentheses. ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, defined at firm level, 5-6 managerial share in wage sum, firm level layer definition. *Independent variables:* see Table 4.

Table B.10: Regression results, ME firm organization, 2012 cross-section, by legal form

Dependent variable	# mgmt. layers, Poisson				Mg. share ∈ [0, 1], GLM			
					Layers		Blossfeld	
GmbH	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.063***		0.018+		0.041+		0.015	
	(0.012)		(0.010)		(0.023)		(0.027)	
Log area		0.043**		0.060***		0.155***		0.107+
		(0.014)		(0.013)		(0.032)		(0.056)
Log sales	0.140***	0.078**						
	(0.012)	(0.025)						
Log # non-mg. employees			0.176***	0.131***				
			(0.011)	(0.026)				
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
# firms	724	215	1,493	452	1,493	452	1,493	452

GmbH & Co. KG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.013*		0.021***		0.061***		0.036*	
	(0.005)		(0.004)		(0.010)		(0.014)	
Log area		0.030***		0.029***		0.077***		0.069***
		(0.006)		(0.005)		(0.013)		(0.018)
Log sales	0.139***	0.118***						
	(0.005)	(0.007)						
Log # non-mg. employees			0.144***	0.117***				
			(0.005)	(0.007)				
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
# firms	3,979	1,212	7,214	2,018	7,214	2,018	7,214	2,018

AG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	−0.002		−0.004		0.041		−0.039	
	(0.010)		(0.010)		(0.033)		(0.059)	
Log area		0.001		0.011		0.038		0.071
		(0.013)		(0.011)		(0.053)		(0.079)
Log sales	0.026***	0.027*						
	(0.008)	(0.010)						
Log # non-mg. employees			0.047***	0.035***				
			(0.008)	(0.010)				
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
# firms	397	228	549	291	549	291	549	291

The tables present the coefficients separately for firms with the legal form *GmbH*, *GmbH & Co. KG* and *AG*. Constant included. Robust standard errors in parentheses. $^{+}$ $p < 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with at least two establishments. *Dependent variable:* (1)-(4) number of managerial layers, (5),(6) managerial share in wage sum, defined by layer, (7),(8) managerial share in wage sum, defined by Blossfeld. *Independent variables:* see Table 4.

B.3 Multi-establishment firms reorganize gradually

Table B.11: Transition dynamics of the managerial organization, 2000-2010 panel, SE firms

# 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 single-establishment firms that transition from a number of managerial layers in year t (given in the rows) to a possibly 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.12: Size at transition, 2000-2010 panel, ME firms

(a) firm-level

# layers in $t/t+1$	0	1	2	3	SE	# firms
0	3.4	3.5	3.7		3.3	10,968
1	3.6	3.7	4.0		3.5	20,327
2		4.0	4.4	4.6	4.0	18,696
3			4.7	5.6	4.9	20,206

(b) 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	**3.4**	3.6^{***}						3.3	10,968
HQ 1/ est. 0	3.7	**3.7**	3.8^{**}	3.8^{***}				3.5	9,252
HQ 1/ est. 1	..	3.8^{+}	**3.9**	4.2^{***}		..		3.7	7,006
HQ 2/ est. 0,1		3.9^{***}	4.2^{*}	**4.3**	4.5^{***}	4.4^{***}		4.0	12,144
HQ 2/ est. 2			..	4.6^{***}	**4.8**	4.8	..	4.3	3,254
HQ 3/ est. 0,1,2				4.5^{***}	5.0^{**}	**5.2**	5.8^{***}	4.8	13,374
HQ 3/ est. 3						6.0^{***}	**6.7**	..	4,606

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 possibly different number of layers in year $t+1$ (given in the columns). Panel (b) the average log number of employees of firms that transition from a managerial structure in year t (given in the rows) to a possibly different managerial structure 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). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. .. denotes cells with fewer than 50 observations. Empty cells contain fewer than .5% of firms. Fewer than .5% of firms exit. Unreported results are similar with sales as outcome variable, although the number of observations is considerably lower.

Table B.13: Transition dynamics of the managerial organization, 2000-2010 panel, by number of establishments

(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	8,085
HQ 1/ est. 0	6	**73**	4	6				10	7,082
HQ 1/ est. 1	1	6	**75**	7				4	4,438
HQ 2/ est. 0,1		4	3	**74**	2	6		9	8,460
HQ 2/ est. 2			.	10	**67**	8	.	3	1,511
HQ 3/ est. 0,1,2				5	2	**83**	2	8	8,280
HQ 3/ est. 3					.	11	**82**	2	1,398

(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	2,883
HQ 1/ est. 0	7	**78**	6	6		.		1	2,170
HQ 1/ est. 1	.	6	**75**	8	.	.		.	2,568
HQ 2/ est. 0,1		4	5	**79**	3	6		1	3,684
HQ 2/ est. 2			.	9	**70**	10	.		1,703
HQ 3/ est. 0,1,2				4	3	**87**	4	1	5,094
HQ 3/ est. 3						8	**88**		3,208

The table displays the percentage share of firms that transition from a managerial structure in year t (given in the rows) to a possibly different managerial structure 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. Panel (a) contains firms that consist of headquarters and exactly one establishment in year t. Panel (b) contains firms that consist of headquarters and at least two establishments in year t. 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.

Table B.14: Transition dynamics of the managerial organization, 2000-2010 panel, by median maximum establishment distance

(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	7,166
HQ 1/ est. 0	6	**75**	4	5				9	5,569
HQ 1/ est. 1	1	6	**75**	7	.			3	3,777
HQ 2/ est. 0,1		5	4	**75**	2	5		8	6,016
HQ 2/ est. 2			.	11	**66**	8		2	1,381
HQ 3/ est. 0,1,2				5	2	**83**	2	7	4,941
HQ 3/ est. 3						12	**82**	.	1,121

(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	**84**	5						6	3,802
HQ 1/ est. 0	6	**73**	5	7		1		7	3,683
HQ 1/ est. 1	.	6	**75**	8		1		2	3,229
HQ 2/ est. 0,1		4	4	**76**	3	8		6	6,128
HQ 2/ est. 2			.	9	**70**	9	.	1	1,873
HQ 3/ est. 0,1,2				5	2	**85**	3	4	8,433
HQ 3/ est. 3						8	**88**	.	3,485

The table displays the percentage share of firms that transition from a managerial structure in year t (given in the rows) to a possibly different managerial structure 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. Panel (a) contains firms with all establishments within the median maximum establishment distance of 170 km in year t. Panel (b) contains firms with establishments above the distance of 170 km in year t. 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.

C Model

Assumption 1. *The maximum value of the helping costs θ_{j0} is .5. The predictability of the production process λ, the helping costs θ_{00} and the learning costs c are such that*

$$\lambda\theta_{00} > c.$$

C.1 Single-establishment firm organization

C.1.1 Lagrangian equation and first order conditions

We use equation (6), which is binding in optimum, to substitute for $n^{\ell}_{0,L}$, $L > \ell > 0$.

$$\begin{aligned}
\mathcal{L} &= n^0_{0,L} w_0 \left(1 + c z^0_{0,L}\right) + n^0_{0,L} \sum_{\ell=1}^{L} \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} w_0 \left(1 + c z^{\ell}_{0,L}\right) + w_0 \left(1 + c\bar{z}_{0,L}\right) \\
&\quad + \xi_{0,L} \left(\tilde{q} - n^0_{0,L}\left(1 - e^{-\lambda \bar{z}_{0,L}}\right)\right) + \varphi_{0,L}\left(n^0_{0,L}\theta_{00} e^{-\lambda z^L_{0,L}} - 1\right) \\
&\quad + \bar{\eta}^L_{0,L}(z^L_{0,L} - \bar{z}_{0,L}) + \sum_{\ell=1}^{L-1} \bar{\eta}^{\ell}_{0,L}(z^{\ell-1}_{0,L} - z^{\ell}_{0,L}) - \bar{\eta}^0_{0,L} z^0_{0,L} - \eta^0_{0,L} n^0_{0,L} \\
\frac{\partial \mathcal{L}}{\partial \bar{z}_{0,L}} &= w_0 c - \xi_{0,L} n^0_{0,L} \lambda e^{-\lambda \bar{z}_{0,L}} - \bar{\eta}^L_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial z^L_{0,L}} & \begin{cases} \overset{L=0}{=} n^0_{0,1}\left(w_0 c - \varphi_{0,1}\theta_{00}\lambda e^{-\lambda z^0_{0,1}}\right) + \bar{\eta}^1_{0,1} - \bar{\eta}^0_{0,1} = 0 \\ \overset{L>0}{=} n^0_{0,L}\left(w_0 c\theta_{00} e^{-\lambda z^{L-1}_{0,L}} - \varphi_{0,L}\theta_{00}\lambda e^{-\lambda z^{L-1}_{0,L}}\right) + \bar{\eta}^L_{0,L} - \bar{\eta}^L_{0,L} = 0 \end{cases} \\
\frac{\partial \mathcal{L}}{\partial z^{\ell}_{0,L}} &= n^0_{0,L} w_0 \left(c\theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} - \lambda\theta_{00} e^{-\lambda z^{\ell}_{0,L}}(1 + c z^{\ell+1}_{0,L})\right) - \bar{\eta}^{\ell}_{0,L} + \bar{\eta}^{\ell+1}_{0,L} = 0 \\
&\text{for } L > \ell > 0,\ L > 1 \\
\frac{\partial \mathcal{L}}{\partial z^0_{0,L}} &\overset{L\geq 0}{=} n^0_{0,L} w_0 \left(c - \lambda\theta_{00} e^{-\lambda z^0_{0,L}}(1 + c z^1_{0,L})\right) + \bar{\eta}^1_{0,L} - \bar{\eta}^0_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial n^0_{0,L}} &= w_0 \left(1 + c z^0_{0,L} + \sum_{\ell=1}^{L} \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} \left(1 + c z^{\ell}_{0,L}\right)\right) \\
&\quad - \xi_{0,L}\left(1 - e^{-\lambda \bar{z}_{0,L}}\right) + \varphi_{0,L}\theta_{00} e^{-\lambda z^L_{0,L}} - \eta^0_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial \xi_{0,L}} &= \tilde{q} - n^0_{0,L}\left(1 - e^{-\lambda \bar{z}_{0,L}}\right) = 0 \\
\frac{\partial \mathcal{L}}{\partial \varphi_{0,L}} &= n^0_{0,L}\theta_{00} e^{-\lambda z^L_{0,L}} - 1 = 0
\end{aligned}$$

C.1.2 Proposition 1: Comparative statics

The second order conditions are, where we substitute $\frac{d\mathcal{L}}{dz^{\ell}_{0,L}}$, $\ell < L$, into $\frac{d^2\mathcal{L}}{dn^0_{0,L} d\tilde{q}}$:

$$\begin{aligned}
\frac{d^2\mathcal{L}}{d\bar{z}_{0,L} d\tilde{q}} &= -\frac{d\xi_{0,L}}{d\tilde{q}} n^0_{0,L}\lambda e^{-\lambda \bar{z}_{0,L}} - \xi_{0,L}\frac{dn^0_{0,L}}{d\tilde{q}}\lambda e^{-\lambda \bar{z}_{0,L}} + \xi_{0,L} n^0_{0,L}\lambda^2 e^{-\lambda \bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}} = 0 \\
\frac{d^2\mathcal{L}}{dz^L_{0,L} d\tilde{q}} & \begin{cases} \overset{L=0}{=} -\frac{d\varphi_{0,1}}{d\tilde{q}}\theta_{00}\lambda e^{-\lambda z^0_{0,1}} + \varphi_{0,1}\theta_{00}\lambda^2 e^{-\lambda z^0_{0,1}}\frac{dz^0_{0,1}}{d\tilde{q}} = 0 \\ \overset{L\geq 0}{=} -w_0 c\lambda e^{-\lambda z^{L-1}_{0,L}}\frac{dz^{L-1}_{0,L}}{d\tilde{q}} - \frac{d\varphi_{0,L}}{d\tilde{q}}\lambda e^{-\lambda z^L_{0,L}} + \varphi_{0,L}\lambda^2 e^{-\lambda z^L_{0,L}}\frac{dz^L_{0,L}}{d\tilde{q}} = 0 \end{cases}
\end{aligned}$$

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

for $L > \ell > 0$, $L > 1$

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

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

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

$$\frac{\partial^2\mathcal{L}}{\partial\varphi_{0,L}\partial\tilde{q}} = \frac{dn^0_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^L_{0,L}} - n^0_{0,L}\theta_{00}\lambda e^{-\lambda z^L_{0,L}}\frac{dz^L_{0,L}}{d\tilde{q}} = 0$$

To show (a): The knowledge of the CEO $\bar{z}_{0,L}$ increases with total output $\tilde{q}$.

1. From $\frac{d^2\mathcal{L}}{d\varphi_{0,L}d\tilde{q}}$:
$$\frac{dz^L_{0,L}}{d\tilde{q}} = \frac{1}{\lambda n^0_{0,L}}\frac{dn^0_{0,L}}{d\tilde{q}}$$

2. From $\frac{d^2\mathcal{L}}{d\xi_{0,L}d\tilde{q}}$:
$$\frac{dn^0_{0,L}}{d\tilde{q}} = \frac{1 - n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}}{1-e^{-\lambda\bar{z}_{0,L}}}$$

3. From $\frac{d^2\mathcal{L}}{dn^0_{0,L}d\tilde{q}}$:
$$\frac{d\xi_{0,L}}{d\tilde{q}} = \frac{\frac{d\varphi_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^L_{0,L}} - \xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}\frac{d\bar{z}_{0,L}}{d\tilde{q}}}{1-e^{-\lambda\bar{z}_{0,L}}}$$

4. From $\frac{d^2\mathcal{L}}{dz^L_{0,L}d\tilde{q}}$, with $\frac{d^2\mathcal{L}}{dz^\ell_{0,L}d\tilde{q}}$, $\ell < L$:
$$\frac{d\varphi_{0,1}}{d\tilde{q}} = \varphi_{0,1}\lambda\frac{dz^0_{0,1}}{d\tilde{q}} \equiv \varphi_{0,1}\lambda f_1(\varphi_{0,1})\frac{dz^L_{0,L}}{d\tilde{q}}$$
$$\frac{d\varphi_{0,2}}{d\tilde{q}} = \varphi_{0,2}\lambda\frac{dz^1_{0,2}}{d\tilde{q}}(1-\theta_{00}e^{-\lambda z^0_{0,2}}) \equiv \varphi_{0,2}\lambda f_2(\varphi_{0,2})\frac{dz^L_{0,L}}{d\tilde{q}}$$
$$\frac{d\varphi_{0,3}}{d\tilde{q}} = \varphi_{0,3}\lambda\frac{dz^2_{0,3}}{d\tilde{q}}\left(1-\frac{\frac{dz^1_{0,3}}{d\tilde{q}}}{\frac{dz^2_{0,3}}{d\tilde{q}}}\right) \equiv \varphi_{0,3}\lambda f_3(\varphi_{0,3})\frac{dz^L_{0,L}}{d\tilde{q}}$$

$f_L(\varphi_{0,L}) > 0$: $f_1(\varphi_{0,1}) = 1 > 0$; $f_2(\varphi_{0,2}) = 1-\theta_{00}e^{-\lambda z^0_{0,2}} > 0$; $f_3(\varphi_{0,3}) = 1 - \frac{dz^1_{0,3}}{d\tilde{q}}\Big/\frac{dz^2_{0,3}}{d\tilde{q}} > 0$ if $\frac{dz^2_{0,3}}{d\tilde{q}} > \frac{dz^1_{0,3}}{d\tilde{q}}$. This is the case if $e^{-\lambda z^1_{0,3}} < e^{-\lambda z^0_{0,3}}(1-\theta_{00}e^{-\lambda z^0_{0,3}})$ $(*)$. $(*)$ holds at $z^2_{0,3} = z^1_{0,3}$, as we can rewrite $(*)$ as $\theta_{00}e^{-\lambda z^0_{0,3}} < 1-\theta_{00}e^{-\lambda z^0_{0,3}}$ for $z^2_{0,3} = z^1_{0,3}$. Both sides of $(*)$ decrease in $z^0_{0,3}$. By $-\lambda e^{-\lambda z^0_{0,3}} < -\lambda e^{-\lambda z^0_{0,3}}(1-2\theta_{00}e^{-\lambda z^0_{0,3}})$, the left-hand side of $(*)$ decreases at a faster rate than the right hand-side of $(*)$ at $z^2_{0,3} = z^1_{0,3}$, so $(*)$ holds for $z^2_{0,3} > z^1_{0,3}$.

5. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,L}d\tilde{q}}$ yields:

$$\frac{d\bar{z}_{0,L}}{d\tilde{q}} = \frac{1}{n^0_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}}} \frac{\xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}} + \frac{\lambda e^{-\lambda\bar{z}_{0,L}}}{1-e^{-\lambda\bar{z}_{0,L}}}\theta_{00}e^{-\lambda z^L_{0,L}}\varphi_{0,L}f_L\left(\varphi_{0,L}\right)}{\xi_{0,L}\lambda e^{-\lambda\bar{z}_{0,L}} + \frac{\lambda e^{-\lambda\bar{z}_{0,L}}}{1-e^{-\lambda\bar{z}_{0,L}}}\theta_{00}e^{-\lambda z^L_{0,L}}\varphi_{0,L}f_L\left(\varphi_{0,L}\right) + \lambda\xi_{0,L}} > 0. \quad \square$$

To show (a): The number $n^\ell_{0,L}$ and the knowledge $z^\ell_{0,L}$ of employees at all below-CEO layers $\ell \leq L$ increase with total output $\tilde{q}$.
Number of employees:

1. $\ell = 0$: $\frac{dn^0_{0,L}}{d\tilde{q}} > 0$ by $\frac{d\bar{z}_{0,L}}{d\tilde{q}} < \frac{1}{\lambda n^0_{0,L}e^{-\lambda\bar{z}_{0,L}}}$.
2. $\ell = L,\ L > 0$: $\frac{dn^L_{0,L}}{d\tilde{q}} = \frac{dn^0_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^{L-2}_{0,L}}f_L(\varphi_{0,L}) > 0$ by $f_L(\varphi_{0,L}) > 0$.
3. $\ell = 1,\ L = 2$: $\frac{dn^1_{0,3}}{d\tilde{q}} > 0$ by $1 > \theta_{00}(e^{-\lambda z^1_{0,3}} + e^{-\lambda z^0_{0,3}})$. $\square$

Knowledge of employees:

1. $\ell = L$: $\frac{dz^L_{0,L}}{d\tilde{q}} = \frac{1}{\lambda n^0_{0,L}}\frac{dn^0_{0,L}}{d\tilde{q}} > 0$ by $\frac{dn^0_{0,L}}{d\tilde{q}} > 0$.
2. $\ell = 0,\ L > 0$: $\frac{dz^0_{0,L}}{d\tilde{q}} = \theta_{00}e^{-\lambda z^0_{0,L}}\frac{dz^1_{0,L}}{d\tilde{q}} > 0$ by $\frac{dz^1_{0,L}}{d\tilde{q}} > 0$.
3. $\ell = 1,\ L = 2$: $\frac{dz^1_{0,3}}{d\tilde{q}} = \frac{dz^2_{0,3}}{d\tilde{q}}\frac{e^{-\lambda z^1_{0,3}}}{e^{-\lambda z^0_{0,3}}(1-\theta_{00}e^{-\lambda z^0_{0,3}})} > 0$ by $\frac{dz^2_{0,3}}{d\tilde{q}} = \frac{dz^L_{0,L}}{d\tilde{q}} > 0$. $\square$

$\frac{dz^{\ell-1}_{0,L}}{d\tilde{q}} < \frac{dz^\ell_{0,L}}{d\tilde{q}}$ by $\theta_{00}e^{-\lambda z^0_{0,L}} < 1$ for $\ell = 1$, $e^{-\lambda z^1_{0,3}} < e^{-\lambda z^0_{0,3}}(1 - \theta_{00}e^{-\lambda z^0_{0,3}})$ for $\ell = 2$ if $z^2_{0,3} \geq z^1_{0,3}$.

To show (a): The marginal benefit of CEO time $\varphi_{0,L}$ increases with total output $\tilde{q}$.
Follows from $\frac{d\varphi_{0,L}}{d\tilde{q}} = \varphi_{0,L}\lambda f_L\left(\varphi_{0,L}\right)\frac{dz^L_{0,L}}{d\tilde{q}} > 0$ by $f_L\left(\varphi_{0,L}\right) > 0$ and $\frac{dz^L_{0,L}}{d\tilde{q}} > 0$. $\square$

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

To show (b): The average cost function $AC_{0,L}(\tilde{q})$ is 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 $\tilde{q} \to \infty$.

$$\begin{aligned} AC_{0,L}(\tilde{q}) &= \frac{C_{0,L}(\tilde{q})}{\tilde{q}} \\ \Rightarrow \quad \frac{dAC_{0,L}(\tilde{q})}{d\tilde{q}} &= \frac{1}{\tilde{q}}\left(\xi_{0,L} - AC_{0,L}\right) = 0 \text{ if } \xi_{0,L} = AC_{0,L} \\ \frac{d^2AC_{0,L}(\tilde{q})}{d\tilde{q}^2} &= -\frac{2}{\tilde{q}^2}\left(\xi_{0,L} - AC_{0,L}\right) + \frac{1}{\tilde{q}}\frac{d\xi_{0,L}}{d\tilde{q}} = \frac{1}{\tilde{q}}\frac{d\xi_{0,L}}{d\tilde{q}} \text{ at } \xi_{0,L} = AC_{0,L} \end{aligned}$$

$\frac{d\xi_{0,L}}{d\tilde{q}} > 0$ if $\varphi_{0,L}f_L(\varphi_{0,L})\theta_{00}e^{-\lambda z^L_{0,L}} > \xi_{0,L}e^{-\lambda\bar{z}_{0,L}}$.

- For $L = 0$, this condition holds $\forall\tilde{q}$.
- For $L > 0$, the condition is equivalent to $e^{\lambda(z^L_{0,L} - z^{L-1}_{0,L})} > (f_L(\varphi_{0,L}))^{-1}$. This condition holds for sufficiently high $\tilde{q}$; in particular, it holds at the MES.

$$\lim_{\tilde{q}\to 0} AC_{0,L}(\tilde{q}) = \infty \text{ because } C_{0,L}(\tilde{q}) \geq w_0 \text{ and } C_{0,L}(\tilde{q}) < \infty \text{ for } \tilde{q} \to 0$$
$$\lim_{\tilde{q}\to \infty} AC_{0,L}(\tilde{q}) = \infty \text{ because } \lim_{\tilde{q}\to\infty} \xi_{0,L} = \infty \text{ by l'Hôpital's rule}$$

C.1.3 The optimal number of layers

We follow Caliendo and Rossi-Hansberg (2012, p. 1454 et seqq.) and 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. 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).

We show that there exists a unique cut-point of the first-order conditions (FOCs) and the respective condition for the minimum efficient scale (MES). We focus on positive solutions for the knowledge levels.
The FOCs (9) and (10) define the optimal knowledge levels recursively:

$$\begin{aligned}
\lambda \bar{z}_{0,L} - \lambda z_{0,L}^L &= \ln\left(\lambda z_{0,L}^0 + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{L} e^{-\lambda z_{0,L}^\ell}\right) - \ln\theta_{00} \\
\lambda z_{0,L}^1 - \lambda z_{0,L}^0 &= \ln\left(\lambda z_{0,L}^2 + \frac{\lambda}{c}\right) && \text{for } L > 2 \\
\lambda z_{0,L}^0 &= \ln\left(\lambda z_{0,L}^1 + \frac{\lambda}{c}\right) + \ln\theta_{00} && \text{for } L > 1
\end{aligned}$$

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

$$\begin{aligned}
\lambda z_{0,1}^0 &= \ln\left(\lambda \bar{z}_{0,1} + \frac{\lambda}{c}\right) + \ln\theta_{00} && \text{for } L = 1 \\
\lambda z_{0,L}^L - \lambda z_{0,L}^{L-1} &= \ln\left(\lambda \bar{z}_{0,L} + \frac{\lambda}{c}\right) && \text{for } L > 1
\end{aligned}$$

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

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

The conditions for the MES have a positive intercept:

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

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

$$\begin{aligned}
\text{MES}: \quad & \frac{dz_{0,L}^L}{d\bar{z}_{0,L}} = \frac{1}{\lambda \bar{z}_{0,L} + \frac{\lambda}{c}} \frac{1}{f_L(\varphi_{0,L})} && > 0 \\
\text{FOC},\ L = 0: \quad & \frac{dz_{0,1}^0}{d\bar{z}_{0,1}} = \frac{\lambda z_{0,1}^0 + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z_{0,1}^0}}{\lambda z_{0,1}^0 + \frac{\lambda}{c} + 1 + f_1(\varphi_{0,L})} && > 0 \\
\text{FOC},\ L = 1: \quad & \frac{dz_{0,2}^1}{d\bar{z}_{0,2}} = \frac{\lambda z_{0,2}^0 + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z_{0,2}^0} + \theta_{00} e^{-\lambda z_{0,2}^1}}{\lambda z_{0,2}^0 + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z_{0,2}^0} + \theta_{00} e^{-\lambda z_{0,2}^0} f_2(\varphi_{0,L})} && > 0
\end{aligned}$$

$$\text{FOC, } L=2: \quad \frac{dz_{0,3}^2}{d\bar{z}_{0,3}} = \frac{\lambda z_{0,3}^0 + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z_{0,3}^{\ell}} + \theta_{00} e^{-\lambda z_{0,3}^{2}}}{\lambda z_{0,3}^0 + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z_{0,3}^{\ell}} + \theta_{00} e^{-\lambda z_{0,3}^{1}} f_3(\varphi_{0,L})} > 0$$

where $f_L(\varphi_{0,L})$ is defined in section C.1.2 (a).

The slope of the conditions for the MES decreases continuously with $\bar{z}_{0,L}$ from a value smaller than 1 with $\lim_{\bar{z}_{0,L}\to\infty} \frac{dz_{0,L}^L}{d\bar{z}_{0,L}} = 0$. The slope of the FOCs is close to 1 with $\lim_{\bar{z}_{0,L}\to\infty} \frac{dz_{0,L}^L}{d\bar{z}_{0,L}} = 1$. 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 layers cannot exceed those of an organization with $L-1$ layers, i.e. $MAC_{0,L-1} \geq MAC_{0,L}$.

C.2 Multi-establishment firm organization

C.2.1 Lagrangian equation and first order conditions

Firm-level: CEO knowledge, allocation of CEO time and output (global product market)

$$\mathcal{L} = \sum_{j=0}^{1} C_{j,\omega}(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}) + (1 - s_{0,\omega} - s_{1,\omega}) w_0 (1 + c\bar{z}_{0,\omega})$$
$$+ \bar{\kappa}_{0,\omega}\left(\sum_{j=0}^{1} s_{j,\omega} - 1\right) - \sum_{j=0}^{1} \kappa_{j,\omega} s_{j,\omega} + \bar{\phi}_{0,\omega}\left(\tilde{q} - \sum_{j=0}^{1} q_{j,\omega}\right) - \sum_{j=0}^{1} \phi_{j,\omega} q_{j,\omega} - \eta_{0,\omega}\bar{z}_{0,\omega}$$

$$\frac{\partial \mathcal{L}}{\partial q_{j,\omega}} = \frac{\partial C_{j,\omega}}{\partial q_{j,\omega}} - \bar{\phi}_{0,\omega} - \phi_{j,\omega} = 0$$
$$\frac{\partial \mathcal{L}}{\partial s_{j,\omega}} = \frac{\partial C_{j,\omega}}{\partial s_{j,\omega}} - w_0(1 + c\bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} - \kappa_{j,\omega} = 0$$
$$\frac{\partial \mathcal{L}}{\partial \bar{z}_{0,\omega}} = \textstyle\sum_{j=0}^{1} \frac{\partial C_{j,\omega}}{\partial \bar{z}_{0,\omega}} + w_0 c(1 - s_{0,\omega} - s_{1,\omega}) - \eta_{0,\omega} = 0$$
$$\frac{\partial \mathcal{L}}{\partial \bar{\kappa}_{0,\omega}} = \textstyle\sum_{j=0}^{1} s_{j,\omega} - 1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \bar{\phi}_{0,\omega}} = \tilde{q} - \textstyle\sum_{j=0}^{1} q_{j,\omega} = 0$$

Establishment-level: Number and knowledge of employees. We use equation (18), which is binding in optimum, to substitute for $n_{j,L}^{\ell}$, $\ell > 0$.

$$\mathcal{L} = n_{j,\omega}^0 w_j \left(1 + c z_{j,\omega}^0\right) + n_{j,\omega}^0 \sum_{\ell=1}^{L_j} \theta_{jj} e^{-\lambda z_{j,\omega}^{\ell-1}} w_j \left(1 + c z_{j,\omega}^{\ell}\right) + s_{j,\omega} w_0 \left(1 + c\bar{z}_{0,\omega}\right)$$
$$+ \xi_{j,\omega}\left(q_{j,\omega} - n_{j,\omega}^0\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right)\right) + \varphi_{j,\omega}\left(n_{j,\omega}^0 \theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} - s_{j,\omega}\right)$$
$$+ \bar{\eta}_{j,\omega}^{L}(z_{j,\omega}^{L_j} - \bar{z}_{0,\omega}) + \sum_{\ell=1}^{L_j} \bar{\eta}_{j,\omega}^{\ell}(z_{j,\omega}^{\ell-1} - z_{j,\omega}^{\ell}) - \bar{\eta}_{j,\omega}^0 z_{j,\omega}^0 - \eta_{j,\omega}^0 n_{j,\omega}^0$$

$$\frac{\partial \mathcal{L}}{\partial z_{j,\omega}^{L_j}} \begin{cases} \overset{L_j=0}{=} n_{j,\omega}^0\left(w_j c - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z_{j,\omega}^0}\right) + \bar{\eta}_{j,\omega}^{L} - \bar{\eta}_{j,\omega}^0 = 0 \\ \overset{L_j>0}{=} n_{j,\omega}^0\left(w_j c\theta_{jj} e^{-\lambda z_{j,\omega}^{L_j-1}} - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z_{j,\omega}^{L_j}}\right) + \bar{\eta}_{j,\omega}^{L} - \bar{\eta}_{j,\omega}^{L_j} = 0 \end{cases}$$

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

for $0 < \ell < L_j - 1$, $L_j > 1$

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

$$\frac{\partial \mathcal{L}}{\partial n^{0}_{j,\omega}} = w_j \left(1 + c z^0_{j,\omega} + \sum_{\ell=1}^{L_j} \theta_{jj} e^{-\lambda z^{\ell-1}_{j,\omega}} \left(1 + c z^{\ell}_{j,\omega} \right) \right)$$

$$- \xi_{j,\omega} \left(1 - e^{-\lambda \bar{z}_{0,\omega}} \right) + \varphi_{j,\omega} \theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - \eta^0_{j,\omega} = 0$$

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

$$\frac{\partial \mathcal{L}}{\partial \varphi_{j,\omega}} = n^0_{j,\omega} \theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - s_{j,\omega} = 0$$

Endogenous variables:

$$e^{\lambda z^{L_j}_{j,\omega}} = \frac{q_{j,\omega}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{\theta_{j0}}{s_{j,\omega}}$$

$$e^{\lambda \left(z^{\ell-1}_{j,\omega} - z^{\ell-2}_{j,\omega} \right)} = \left(1 + c z^{\ell}_{j,\omega} \right) \frac{\lambda}{c} \quad \forall \ell = 2, ..., L_j$$

$$e^{\lambda z^{0}_{j,\omega}} = \left(1 + c z^{1}_{j,\omega} \right) \frac{\lambda}{c} \theta_{jj} \quad \text{for } L_j > 0$$

$$\xi_{j,\omega} = \frac{w_j \left(1 + c z^0_{j,\omega} + \frac{c}{\lambda} + \mathbb{1}(L_j \geq 1) \theta_{jj} \frac{c}{\lambda} \sum_{\ell=1}^{L_j} e^{-\lambda z^{\ell-1}_{j,\omega}} \right)}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$$

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

C.2.2 Proposition 2: Allocation of output and CEO time

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

The first order conditions imply:

$\frac{\partial \mathcal{L}}{\partial q_{j,\omega}}$: If $\phi_{j,\omega} = 0 \, \forall j$, i.e. if there is positive production at both establishments,

$$\phi_{0,\omega} = \frac{\partial C}{\partial q_{0,\omega}} - \bar{\phi}_{0,\omega} = \frac{\partial C}{\partial q_{1,\omega}} - \bar{\phi}_{0,\omega} = \phi_{1,\omega} = 0 \quad \text{for } q_{0,\omega}, q_{1,\omega} > 0$$

$$\Rightarrow \quad \frac{\partial C}{\partial q_{0,\omega}} = \xi_{0,\omega} = \xi_{1,\omega} = \frac{\partial C}{\partial q_{1,\omega}}.$$

$\frac{\partial \mathcal{L}}{\partial s_{j,\omega}}$: If $\kappa_{j,\omega} = 0 \forall j$, i.e. if the CEO spends positive time on both establishments,

$$\kappa_{0,\omega} = \frac{\partial C}{\partial s_{0,\omega}} - w_0 (1 + c \bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} =$$

$$\kappa_{1,\omega} = \frac{\partial C}{\partial s_{1,\omega}} - w_0 (1 + c \bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} = 0 \quad \text{for } s_{0,\omega}, s_{1,\omega} > 0$$

$$\Rightarrow \quad \varphi_{0,\omega} = \varphi_{1,\omega}.$$

Corollary 1. To show: It is not optimal to produce at two establishments with the same number of below-CEO management layers $L_0 = L_1$ if $\theta_{10} > \theta_{00}$ and $w_1 \geq w_0$.

Both the marginal production costs $\xi_{j,\omega}$ and the marginal benefit of CEO time $\varphi_{j,\omega}$ are equal across establishments if the firm produces at two establishments. $\varphi_{j,\omega}$ is a function of θ_{10}, but $\xi_{j,\omega}$ is not. Production at two establishments with $L_0 = L_1$ and symmetric helping costs $\theta_{10} = \theta_{00}$ but wages $w_1 \geq w_0$ therefore violates Proposition 2.
To see this, consider $L_j = 0\,\forall j$. The following equations cannot be simultaneously fulfilled:

$$w_0(1 + cz^0_{0,\omega}) = w_1(1 + cz^0_{1,\omega}) \quad (\text{from } \xi_{0,\omega} = \xi_{1,\omega})$$
$$w_0\theta_{10}e^{\lambda z^0_{0,\omega}} = w_1\theta_{00}e^{\lambda z^0_{1,\omega}} \quad (\text{from } \varphi_{0,\omega} = \varphi_{1,\omega})$$

For $w_0 = w_1$, $\xi_{0,\omega} = \xi_{1,\omega}$ requires that the knowledge levels are the same, but $\varphi_{0,\omega} = \varphi_{1,\omega}$ requires that they are different. For $w_0 < w_1$, both equations define $z^0_{1,\omega}$ as increasing function of $z^0_{0,\omega}$. They do not intersect, because both the intercept and the slope of $\xi_{0,\omega} = \xi_{1,\omega}$ are smaller than the ones of $\varphi_{0,\omega} = \varphi_{1,\omega}$. Hence, there is no pair of $z^0_{0,\omega}$, $z^0_{1,\omega}$ that fulfills both equations.

C.2.3 Proposition 3: Comparative statics with respect to $\tilde{q}$

The second order conditions for $q_{j,\omega} > 0$, $s_{j,\omega} > 0\,\forall j$ are, substituting $\frac{d\mathcal{L}}{dz^\ell_{j,\omega}}$, $\ell \leq L_j$, into $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}}$:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}} = \frac{d\xi_{0,\omega}}{d\tilde{q}} - \frac{d\xi_{1,\omega}}{d\tilde{q}} = 0$$

$$\frac{d^2\mathcal{L}}{ds_{0,\omega}d\tilde{q}} - \frac{d^2\mathcal{L}}{ds_{1,\omega}d\tilde{q}} = \frac{d\varphi_{0,\omega}}{d\tilde{q}} - \frac{d\varphi_{1,\omega}}{d\tilde{q}} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}} = -\sum_{j=0}^{1}\frac{d\xi_{j,\omega}}{d\tilde{q}}n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}} - \sum_{j=0}^{1}\xi_{j,\omega}\frac{dn^0_{j,\omega}}{d\tilde{q}}\lambda e^{-\lambda\bar{z}_{0,\omega}} + \sum_{j=0}^{1}\xi_{j,\omega}n^0_{j,\omega}\lambda^2e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} = 0$$

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

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

$$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\tilde{q}} \begin{cases} \overset{L_j=0}{=} -\frac{d\varphi_{j,\omega}}{d\tilde{q}}\theta_{j0}\lambda e^{-\lambda z^0_{j,\omega}} + \varphi_{j,\omega}\theta_{j0}\lambda^2e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\tilde{q}} = 0 \\ \overset{L_j>0}{=} -w_jc\theta_{jj}\lambda e^{-\lambda z^{L_j-1}_{j,\omega}}\frac{dz^{L_j-1}_{j,\omega}}{d\tilde{q}} - \frac{d\varphi_{j,\omega}}{d\tilde{q}}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}} + \varphi_{j,\omega}\theta_{j0}\lambda^2e^{-\lambda z^{L_j}_{j,\omega}}\frac{dz^{L_j}_{j,\omega}}{d\tilde{q}} = 0 \end{cases}$$

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

for $0 < \ell < L_j$, $L_j > 1$

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\tilde{q}} \overset{L_j>0}{=} \lambda^2\theta_{jj}e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\tilde{q}}(1 + cz^1_{j,\omega}) - \lambda\theta_{jj}e^{-\lambda z^0_{j,\omega}}c\frac{dz^1_{j,\omega}}{d\tilde{q}} = 0 \quad \text{for } L_j > 0$$

$$\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}} = -\frac{d\xi_{j,\omega}}{d\tilde{q}}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - \xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} + \frac{d\varphi_{j,\omega}}{d\tilde{q}}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}} = \frac{dq_{j,\omega}}{d\tilde{q}} - \frac{dn^0_{j,\omega}}{d\tilde{q}}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} = 0$$

$$\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\tilde{q}} = \frac{dn^0_{j,\omega}}{d\tilde{q}}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}} - n^0_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}}\frac{dz^{L_j}_{j,\omega}}{d\tilde{q}} - \frac{ds_{j,\omega}}{d\tilde{q}} = 0$$

To show (a): CEO knowledge $\bar{z}_{0,\omega}$ increases with total output $\tilde{q}$.

1. As will be shown below, $\frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0$ and $\frac{dz^{\ell}_{j,\omega}}{d\tilde{q}} = 0$, $\ell = 0, ..., L_j$. $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}}$ yields:

$$\frac{d\xi_{j,\omega}}{d\tilde{q}} = -\frac{\xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}}$$

2. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}}$:

$$\frac{dn^0_{j,\omega}}{d\tilde{q}} = \frac{\frac{dq_{j,\omega}}{d\tilde{q}} - n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}$$

3. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}}$ together with $\frac{d\xi_{0,\omega}}{d\tilde{q}} - \frac{d\xi_{1,\omega}}{d\tilde{q}} = 0$ and $1 - \frac{dq_{0,\omega}}{d\tilde{q}} - \frac{dq_{1,\omega}}{d\tilde{q}} = 0$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} = \frac{1-e^{-\lambda\bar{z}_{0,\omega}}}{\lambda\tilde{q}(1+e^{-\lambda\bar{z}_{0,\omega}})} > 0$$

To show (a): The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n^{\ell}_{j,\omega}$, $\ell = 0, ..., L_j$, increases with total output $\tilde{q}$.

- $\ell = 0$: Follows from $\sum_{j=0}^{1}\frac{dn^0_{j,\omega}}{d\tilde{q}} = \frac{1-\sum_{j=0}^{1} n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}$ and $\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} < \frac{1-e^{-\lambda\bar{z}_{0,\omega}}}{\lambda\tilde{q}e^{-\lambda\bar{z}_{0,\omega}}}$.
- $\ell > 0$: Follows from $\sum_{j=0}^{1}\frac{dn^0_{j,\omega}}{d\tilde{q}} > 0$ and $\frac{dz^{\ell}_{j,\omega}}{d\tilde{q}} = 0$ (see below).

To show (a): The marginal benefit of CEO time $\varphi_{j,\omega}$ does not vary with total output $\tilde{q}$.
From $\frac{d^2\mathcal{L}}{dq_{j,\omega}d\tilde{q}}$ and $\frac{d^2\mathcal{L}}{ds_{j,\omega}d\tilde{q}}$, we know:

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}} = \frac{d\varphi_{1,\omega}}{d\tilde{q}} \quad \text{and} \quad \frac{d\xi_{0,\omega}}{d\tilde{q}} = \frac{d\xi_{1,\omega}}{d\tilde{q}}.$$

Substituting for $\frac{d\xi_{j,\omega}}{d\tilde{q}}$ from $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}}$ implies, with $\xi_{0,\omega} = \xi_{1,\omega}$:

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} = \frac{d\varphi_{1,\omega}}{d\tilde{q}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}$$

This equation holds if $\frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0$, or $\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} = \theta_{10}e^{-\lambda z^{L_1}_{1,\omega}}$, which is at odds with $\varphi_{0,\omega} = \varphi_{1,\omega}$.

To show (a): The knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\ell = 0, ..., L_j$, is constant.
$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\tilde{q}}$ implies:

$$\begin{aligned}
&\frac{d\varphi_{j,\omega}}{d\tilde{q}} = \varphi_{j,\omega}\lambda\frac{dz^0_{j,\omega}}{d\tilde{q}} \text{ if } L_j = 0\\
&\frac{d\varphi_{j,\omega}}{d\tilde{q}} = \varphi_{j,\omega}\lambda\frac{dz^{L_j}_{j,\omega}}{d\tilde{q}} - w_j c\frac{\theta_{jj}}{\theta_{j0}}e^{\lambda(z^{L_j}_{j,\omega}-z^{L_j-1}_{j,\omega})}\frac{dz^{L_j-1}_{j,\omega}}{d\tilde{q}} \text{ if } L_j > 0\\
&\Rightarrow \quad \frac{dz^{\ell}_{j,\omega}}{d\tilde{q}} = 0\,\forall j,\ell \text{ by } \frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0\,\forall j
\end{aligned}$$

To show (b): The share of CEO time $s_{j,\omega}$ and the number of employees at all below-CEO layers $n^{\ell}_{j,\omega}$, $\forall \ell < L$, increase with total output $\tilde{q}$ at the establishment and decrease at the headquarters, unless $L_0 > L_1$ and wages w_1 are too high. Local output $q_{j,\omega}$ increases at the establishment.

1. Combining $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}}$, $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\tilde{q}}$, $\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}}$, and $\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\tilde{q}}$ yields:

$$\frac{ds_{0,\omega}}{d\tilde{q}} = \frac{1}{(1-e^{-\lambda\bar{z}_{0,\omega}})(1+e^{-\lambda\bar{z}_{0,\omega}})} \frac{\theta_{00}\theta_{10}}{\theta_{10}e^{\lambda z^{L_0}_{0,\omega}} - \theta_{00}e^{\lambda z^{L_1}_{1,\omega}}}; \quad \frac{ds_{1,\omega}}{d\tilde{q}} = -\frac{ds_{0,\omega}}{d\tilde{q}}$$

   The sign of $ds_{0,\omega}/d\tilde{q}$ is determined by the second factor.

   - For $L_0 = L_1 \in \{0,1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{10}e^{\lambda z^{L_0}_{0,\omega}} - \theta_{00}e^{\lambda z^{L_1}_{1,\omega}} < 0$ if $w_0 > w_1$. $w_0 > w_1$ because we assume $w_0 \geq w_1$ in section 4.1; $q_j > 0\,\forall j$ is not optimal for a symmetric number of layers with $w_0 = w_1$ and $\theta_{10} > \theta_{00}$ by Corollary 1.
   - For $L_0 < L_1 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{10}e^{\lambda z^{L_0}_{0,\omega}} - \theta_{00}e^{\lambda z^{L_1}_{1,\omega}} < 0$ if $w_0 \geq w_1$.
   - For $L_1 < L_0 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{10}e^{\lambda z^{L_0}_{0,\omega}} - \theta_{00}e^{\lambda z^{L_1}_{1,\omega}} < 0$ 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.

2. By equation (18), $\text{sgn}\left(\frac{dn^{\ell}_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{dn^{0}_{j,\omega}}{d\tilde{q}}\right)$. By $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\tilde{q}}$, $\text{sgn}\left(\frac{dn^{0}_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{ds_{j,\omega}}{d\tilde{q}}\right)$, i.e. the number of employees varies as the share of CEO time.

3. By $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}}$, $\text{sgn}\left(\frac{dq_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{ds_{j,\omega}}{d\tilde{q}}\right)$ for $\frac{ds_{j,\omega}}{d\tilde{q}} > 0$. As $\frac{dq_{0,\omega}}{d\tilde{q}} + \frac{dq_{1,\omega}}{d\tilde{q}} = 1$ and $\frac{ds_{0,\omega}}{d\tilde{q}} + \frac{ds_{1,\omega}}{d\tilde{q}} = 0$, the sign is indeterminate if $\frac{ds_{j,\omega}}{d\tilde{q}} < 0$.

To show (c): The cost function $C_{0,\omega}(\tilde{q})$ strictly increases with total output $\tilde{q}$. Follows from $\frac{\partial C_{0,\omega}(\tilde{q})}{\partial \tilde{q}} = \bar{\phi}_{0,\omega} \geq 0$ with $\bar{\phi}_{0,\omega} = \frac{w_0 c(e^{\lambda\bar{z}_{0,\omega}}-1)}{\lambda\tilde{q}}$.

To show (c): The marginal production cost $\xi_{j,\omega}$ decreases with total output $\tilde{q}$. Follows from $\frac{d\xi_{j,\omega}}{d\tilde{q}} = -\frac{\xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}}$ and $\frac{d\bar{z}_{0,\omega}}{d\tilde{q}} > 0$.

Full symmetry. Under full symmetry, the cost function coincides with the cost function of a single establishment firm, so Proposition 1 applies.

C.2.4 Proposition 4: Comparative statics with respect to θ_{10}

The second order conditions for $q_{j,\omega} > 0$, $s_{j,\omega} > 0\,\forall j$ are, substituting $\frac{d\mathcal{L}}{dz^{\ell}_{j,\omega}}$, $\ell \leq L_j$, into $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\theta_{10}}$:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$$

$$\frac{d^2\mathcal{L}}{ds_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{ds_{1,\omega}d\theta_{10}} = \frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$$

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

$$\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\theta_{10}} = \frac{ds_{0,\omega}}{d\theta_{10}} + \frac{ds_{1,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}} = -\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{dz^{L_0}_{0,\omega}d\theta_{10}} \begin{cases} \overset{L_0=0}{=} -\frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{00}\lambda e^{-\lambda z^0_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^0_{0,\omega}}\frac{dz^0_{0,\omega}}{d\theta_{10}} = 0 \\ \overset{L_0>0}{=} -w_0 c\lambda\theta_{00}e^{-\lambda z^{L_0-1}_{0,\omega}}\frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{0,\omega}}{d\theta_{10}}\lambda\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^{L_0}_{0,\omega}}\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} = 0 \end{cases}$$
$$\frac{d^2\mathcal{L}}{dz^{L_1}_{1,\omega}d\theta_{10}} \begin{cases} \overset{L_1=0}{=} -\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}\lambda e^{-\lambda z^0_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^0_{1,\omega}}\frac{dz^0_{1,\omega}}{d\theta_{10}} - \varphi_{1,\omega}\lambda e^{-\lambda z^0_{1,\omega}} = 0 \\ \overset{L_0>0}{=} -w_1 c\lambda\theta_{11}e^{-\lambda z^{L_1-1}_{1,\omega}}\frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}}\lambda\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} \\ -\varphi_{1,\omega}\lambda e^{-\lambda z^{L_1}_{1,\omega}} = 0 \end{cases}$$
$$\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}d\theta_{10}} = -\lambda c e^{-\lambda z^{\ell-1}_{j,\omega}}\frac{dz^{\ell-1}_{j,\omega}}{d\theta_{10}} + \lambda^2 e^{-\lambda z^{\ell}_{j,\omega}}\frac{dz^{\ell}_{j,\omega}}{d\theta_{10}}(1 + cz^{\ell+1}_{j,\omega}) - \lambda e^{-\lambda z^{\ell}_{j,\omega}}c\frac{dz^{\ell+1}_{j,\omega}}{d\theta_{10}} = 0$$

for $0 < \ell < L_j$, $L_j > 1$

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\theta_{10}} \overset{L_j>0}{=} \lambda^2\theta_{jj}e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\theta_{10}}(1 + cz^1_{j,\omega}) - \lambda\theta_{jj}e^{-\lambda z^0_{j,\omega}}c\frac{dz^1_{j,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{dn^0_{0,\omega}d\theta_{10}} = -\frac{d\xi_{0,\omega}}{d\theta_{10}}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - \xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} = 0$$
$$\frac{d^2\mathcal{L}}{dn^0_{1,\omega}d\theta_{10}} = -\frac{d\xi_{1,\omega}}{d\theta_{10}}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - \xi_{1,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} + \varphi_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}} = 0$$
$$\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}} = \frac{dq_{j,\omega}}{d\theta_{10}} - \frac{dn^0_{j,\omega}}{d\theta_{10}}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\varphi_{0,\omega}d\theta_{10}} = \frac{dn^0_{0,\omega}}{d\theta_{10}}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - n^0_{0,\omega}\theta_{00}\lambda e^{-\lambda z^{L_0}_{0,\omega}}\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} - \frac{ds_{0,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\varphi_{1,\omega}d\theta_{10}} = \frac{dn^0_{1,\omega}}{d\theta_{10}}\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} - n^0_{1,\omega}\theta_{10}\lambda e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} - \frac{ds_{1,\omega}}{d\theta_{10}} + n^0_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}} = 0$$

To show (a): CEO knowledge $\bar{z}_{0,\omega}$ increases with the helping costs θ_{10}.

1. The two equations $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\theta_{10}}$ $j = 0, 1$ yield, together with $\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$, $\xi_{0,\omega} = \xi_{1,\omega}$, $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$ and $\varphi_{0,\omega} = \varphi_{1,\omega}$:

$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}e^{\lambda z^{L_0}_{0,\omega}}}{\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}}}$$

2. Substituting into $\frac{d^2\mathcal{L}}{dn^0_{0,\omega}d\theta_{10}}$ results in:

$$\frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{\theta_{00}\varphi_{0,\omega} - \xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}(\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})}{(1 - e^{-\lambda\bar{z}_{0,\omega}})(\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})}$$

3. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$:

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

4. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\theta_{10}}$, together with $\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$ and $-\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$,

yields:

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

5. Combining the two expressions for $\frac{d\xi_{0,\omega}}{d\theta_{10}}$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}\theta_{00}}{\lambda\xi_{0,\omega}(1+e^{-\lambda\bar{z}_{0,\omega}})(\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})} > 0$$

$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ if $\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}} > 0$.

- For $L_0 = L_1 \in \{0,1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}} > 0$ if $w_0 > w_1$. $w_0 > w_1$ because we assume $w_0 \geq w_1$ in section 4.1; $q_j > 0 \, \forall j$ is not optimal for a symmetric number of layers with $w_0 = w_1$ and $\theta_{10} > \theta_{00}$ by Corollary 1.
- For $L_0 < L_1 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}} > 0$ if $w_0 \geq w_1$.
- For $L_1 < L_0 \in \{1,2\}$, $\varphi_{0,\omega} = \varphi_{1,\omega}$ implies that $\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}} > 0$ 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.

To show (a): The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with the helping costs θ_{10}.
Follows from $\frac{d\varphi_{1,\omega}}{d\theta_{10}} = \frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}e^{\lambda z^{L_0}_{0,\omega}}}{\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}}} > 0$ if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.

To show (a): The knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\forall \ell \leq L_j$ increases with the helping costs θ_{10}.
$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\theta_{10}}$ yields:

$$\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} \overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$$\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} - \frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}} \overset{L_0>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} - \frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}} \overset{L_1>0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

Together with $\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}d\theta_{10}}$, $0 \leq \ell < L_j$, these equations imply that $\frac{dz^{\ell}_{j,\omega}}{d\theta_{10}} > 0$ if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.

To show (a): The increase of the knowledge levels with the helping costs θ_{10} is stronger at higher than at lower layers of an establishment. If both establishments have the same number of layers, the increase is stronger at the establishment than at the headquarters.
Higher vs. lower layers: From $\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}d\theta_{10}}$:

$$\frac{dz^0_{j,\omega}}{d\theta_{10}} = \theta_{jj}e^{-\lambda z^0_{j,\omega}}\frac{dz^1_{j,\omega}}{d\theta_{10}} < \frac{dz^1_{j,\omega}}{d\theta_{10}}; \qquad \frac{dz^1_{j,\omega}}{d\theta_{10}} = \frac{e^{-\lambda z^1_{j,\omega}}}{e^{-\lambda z^0_{j,\omega}}(1-\theta_{jj}e^{-\lambda z^0_{j,\omega}})}\frac{dz^2_{j,\omega}}{d\theta_{10}} < \frac{dz^2_{j,\omega}}{d\theta_{10}}$$

Establishment vs. headquarters: From $\frac{d^2\mathcal{L}}{dz_{j,\omega}^{L_j} d\theta_{10}}$:

$$\frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} \overset{L_j=0\forall j}{=} \frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}; \qquad \frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} - \frac{dz_{1,\omega}^{L_1-1}}{d\theta_{10}} \overset{L_0=L_1>0}{=} \frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} - \frac{dz_{0,\omega}^{L_0-1}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

To show (a): 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 $\frac{d^2\mathcal{L}}{d\xi_{j,\omega} d\theta_{10}}$, with $-\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$:

$$\sum_{j=0}^{1} \frac{dn_{j,\omega}^{0}}{d\theta_{10}} = -\sum_{j=0}^{1} \frac{n_{j,\omega}^{0} \lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} < 0 \quad \text{as} \quad \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$$

$\ell > 0$: Follows from $\sum_{j=0}^{1} \frac{dn_{j,\omega}^{0}}{d\theta_{10}} < 0$ and $\frac{dz_{j,\omega}^{\ell}}{d\theta_{10}} > 0$.

To show (b): The share of CEO time $s_{j,\omega}$, local output $q_{j,\omega}$ and the number of employees at all below-CEO layers $n_{j,\omega}^{\ell}$ increase at the headquarters and decrease at the establishment.

1. $\frac{d^2\mathcal{L}}{d\xi_{j,\omega} d\theta_{10}}$ and $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega} d\theta_{10}}$ for $j = 0, 1$ imply, together with $\sum_{j=0}^{1} \frac{dq_{j,\omega}}{d\theta_{10}} = 0$ and $\sum_{j=0}^{1} \frac{ds_{j,\omega}}{d\theta_{10}} = 0$:

$$\begin{aligned} \frac{dq_{0,\omega}}{d\theta_{10}} &= \frac{e^{\lambda z_{0,\omega}^{L_0}} e^{\lambda z_{1,\omega}^{L_1}}}{\theta_{00} e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10} e^{\lambda z_{0,\omega}^{L_0}}} \left(\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} \right. \\ &\quad \left. + \left(1 - e^{-\lambda \bar{z_{0,\omega}}}\right) \left(\lambda \sum_{j=0}^{1} n_{j,\omega}^{0} \theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} \frac{dz_{j,\omega}^{L_j}}{d\theta_{10}} - n_{1,\omega}^{0} e^{-\lambda z_{1,\omega}^{L_1}} \right) \right) > 0 \quad \text{if} \quad \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0 \\ \Rightarrow \quad \frac{dq_{1,\omega}}{d\theta_{10}} &< 0 \end{aligned}$$

2. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega} d\theta_{10}}$:

$$\begin{aligned} \frac{dn_{j,\omega}^{0}}{d\theta_{10}} &= \frac{\frac{dq_{j,\omega}}{d\theta_{10}} - n_{j,\omega}^{0} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \\ &> 0 \quad \text{for } j = 0 \text{ by } \frac{dq_{0,\omega}}{d\theta_{10}} > n_{0,\omega}^{0} \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} \\ &< 0 \quad \text{for } j = 1 \end{aligned}$$

3. $\frac{d^2\mathcal{L}}{d\varphi_{0,\omega} d\theta_{10}}$ yields (with a little algebra):

$$\frac{ds_{0,\omega}}{d\theta_{10}} > 0 \quad \text{by} \quad \frac{dn_{0,\omega}^{0}}{d\theta_{10}} > n_{0,\omega}^{0} \lambda \frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} \qquad \Rightarrow \qquad \frac{ds_{1,\omega}}{d\theta_{10}} < 0$$

To show (c): The cost function $C_{0,\omega}(\tilde{q})$ and the marginal production cost $\xi_{j,\omega}$ increase with the helping costs θ_{10}.
Follows from the envelope theorem and $\frac{d\xi_{1,\omega}}{d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} \frac{\xi_{0,\omega}\lambda}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$ if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.

C.2.5 Comparative statics with respect to w_1

The second order conditions for $q_{j,\omega} > 0, s_{j,\omega} > 0 \forall j$ are, substituting $\frac{d\mathcal{L}}{dz^{\ell}_{j,\omega}}$, $\ell \leq L_j$, into $\frac{d^2\mathcal{L}}{dn^0_{j,\omega} d\tilde{q}}$:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}dw_1} - \frac{d^2\mathcal{L}}{dq_{1,\omega}dw_1} = \frac{d\xi_{0,\omega}}{dw_1} - \frac{d\xi_{1,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{ds_{0,\omega}dw_1} - \frac{d^2\mathcal{L}}{ds_{1,\omega}dw_1} = \frac{d\varphi_{0,\omega}}{dw_1} - \frac{d\varphi_{1,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}dw_1} = -\sum_{j=0}^{1} \frac{d\xi_{j,\omega}}{dw_1} n^0_{j,\omega} \lambda e^{-\lambda \bar{z}_{0,\omega}} - \sum_{j=0}^{1} \xi_{j,\omega} \frac{dn^0_{j,\omega}}{dw_1} \lambda e^{-\lambda \bar{z}_{0,\omega}} + \sum_{j=0}^{1} \xi_{j,\omega} n^0_{j,\omega} \lambda^2 e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega} w_1} = \frac{ds_{0,\omega}}{dw_1} + \frac{ds_{1,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}dw_1} = -\frac{dq_{0,\omega}}{dw_1} - \frac{dq_{1,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{dz^{L_0}_{0,\omega}dw_1} \begin{cases} \overset{L_0=0}{=} -\frac{d\varphi_{0,\omega}}{dw_1}\theta_{00}\lambda e^{-\lambda z^0_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^0_{0,\omega}} \frac{dz^0_{0,\omega}}{dw_1} = 0 \\ \overset{L_0>0}{=} -w_0 c\theta_{00}\lambda e^{-\lambda z^{L_0-1}_{0,\omega}} \frac{dz^{L_0-1}_{0,\omega}}{dw_1} - \frac{d\varphi_{0,\omega}}{dw_1}\theta_{00}\lambda e^{-\lambda z^{L_0}_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^{L_0}_{0,\omega}} \frac{dz^{L_0}_{0,\omega}}{dw_1} = 0 \end{cases}$$

$$\frac{d^2\mathcal{L}}{dz^{L_1}_{1,\omega}dw_1} \begin{cases} \overset{L_1=0}{=} c - \frac{d\varphi_{1,\omega}}{dw_1}\theta_{10}\lambda e^{-\lambda z^0_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^0_{1,\omega}} \frac{dz^0_{1,\omega}}{dw_1} = 0 \\ \overset{L_1>0}{=} c\theta_{11} e^{-\lambda z^{L_1-1}_{1,\omega}} - w_1 c\theta_{11}\lambda e^{-\lambda z^{L_1-1}_{1,\omega}} \frac{dz^{L_1-1}_{1,\omega}}{dw_1} - \frac{d\varphi_{1,\omega}}{dw_1}\theta_{10}\lambda e^{-\lambda z^{L_1}_{1,\omega}} \\ +\varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^{L_1}_{1,\omega}} \frac{dz^{L_1}_{1,\omega}}{dw_1} = 0 \end{cases}$$

$$\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}dw_1} = -\lambda c e^{-\lambda z^{\ell-1}_{j,\omega}} \frac{dz^{\ell-1}_{j,\omega}}{dw_1} + \lambda^2 e^{-\lambda z^{\ell}_{j,\omega}} \frac{dz^{\ell}_{j,\omega}}{dw_1}(1 + cz^{\ell+1}_{j,\omega}) - \lambda e^{-\lambda z^{\ell}_{j,\omega}} c \frac{dz^{\ell+1}_{j,\omega}}{dw_1} = 0$$

for $0 < \ell < L_j$, $L_j > 1$

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}dw_1} \overset{L_j>0}{=} \lambda^2\theta_{jj} e^{-\lambda z^0_{j,\omega}} \frac{dz^0_{j,\omega}}{dw_1}(1 + cz^1_{j,\omega}) - \lambda\theta_{jj} e^{-\lambda z^0_{j,\omega}} c \frac{dz^1_{j,\omega}}{dw_1} = 0 \quad \text{for } L_j > 0$$

$$\frac{d^2\mathcal{L}}{dn^0_{0,\omega}dw_1} = -\frac{d\xi_{0,\omega}}{dw_1}(1 - e^{-\lambda \bar{z}_{0,\omega}}) - \xi_{0,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} + \frac{d\varphi_{0,\omega}}{dw_1}\theta_{00} e^{-\lambda z^{L_0}_{j0\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{dn^0_{1,\omega}dw_1} = 1 + cz^0_{1,\omega} + \mathbb{1}(L_1 \geq 1) \sum_{\ell=0}^{L_1} \theta_{11} e^{-\lambda z^{\ell-1}_{1,\omega}} (1 + cz^{\ell}_{1,\omega})$$
$$- \frac{d\xi_{1,\omega}}{dw_1}(1 - e^{-\lambda \bar{z}_{0,\omega}}) - \xi_{1,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} + \frac{d\varphi_{1,\omega}}{dw_1}\theta_{10} e^{-\lambda z^{L_1}_{1,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\xi_{j,\omega}dw_1} = \frac{dq_{j,\omega}}{dw_1} - \frac{dn^0_{j,\omega}}{dw_1}(1 - e^{-\lambda \bar{z}_{0,\omega}}) - n^0_{j,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} = 0$$

$$\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}dw_1} = \frac{dn^0_{j,\omega}}{dw_1}\theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - n^0_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}} \frac{dz^{L_j}_{j,\omega}}{dw_1} - \frac{ds_{j,\omega}}{dw_1} = 0$$

Suppose that $L_0 \leq L_1$, or $L_0 > L_1$ and w_1 is sufficiently low.

► CEO knowledge $\bar{z}_{0,\omega}$ increases with local wages w_1.

1. $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}dw_1}$ $j=0,1$ yield, with $\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$, $\xi_{0,\omega} = \xi_{1,\omega}$, and $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$:

$$\frac{d\varphi_{j,\omega}}{dw_1} = \frac{\hat{w}_1 e^{\lambda z^{L_0}_{0,\omega}} e^{\lambda z^{L_1}_{1,\omega}}}{\theta_{00} e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10} e^{\lambda z^{L_0}_{0,\omega}}}$$

   where we define $\hat{w}_1 \equiv 1 + cz^0_{1,\omega} + \mathbb{1}(L_1 \geq 1) \sum_{\ell=0}^{L_1} \theta_{11} e^{-\lambda z^{\ell-1}_{1,\omega}} (1 + cz^{\ell}_{1,\omega})$.

2. Substituting into $\frac{d^2\mathcal{L}}{dn^0_{0,\omega}dw_1}$ yields:

$$\frac{d\xi_{j,\omega}}{dw_1} = \frac{\frac{d\varphi_{j,\omega}}{dw_1}\theta_{00} e^{-\lambda z^{L_0}_{0,\omega}} - \xi_{0,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1}}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$$

3. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}dw_1}$:

$$\frac{dn^0_{j,\omega}}{dw_1} = \frac{\frac{dq_{j,\omega}}{dw_1} - n^0_{j,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1}}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$$

4. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}dw_1}$, with $\frac{d\xi_{0,\omega}}{dw_1} - \frac{d\xi_{1,\omega}}{dw_1} = 0$ and $-\frac{dq_{0,\omega}}{dw_1} - \frac{dq_{1,\omega}}{dw_1} = 0$, yields:

$$\frac{d\bar{z}_{0,\omega}}{dw_1} = \frac{\hat{w}_1 \theta_{00} e^{\lambda z^{L_1}_{1,\omega}}}{\lambda \xi_{0,\omega}(1 + e^{\lambda \bar{z}_{0,\omega}})(\theta_{00} e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10} e^{\lambda z^{L_0}_{0,\omega}})}$$

$\Rightarrow \frac{d\bar{z}_{0,\omega}}{dw_1} > 0$ under the same conditions as $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ (i.e., the conditions stated above)

▶ The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with local wages w_1.

▶ The knowledge of the employees at all below-CEO layers at the headquarters $z^{\ell}_{0,\omega}$, $\forall \ell \leq L_0$ (establishment $z^{\ell}_{1,\omega}$, $\forall \ell \leq L_1$) increases with wages w_1 (if w_1 is sufficiently high).
$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}dw_1}$ yields:

$$\frac{dz^{L_0}_{0,\omega}}{dw_1} \overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{dw_1} \qquad \frac{dz^{L_1}_{1,\omega}}{dw_1} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{dw_1} - \frac{1}{\lambda w_1}$$

$$\frac{dz^{L_0}_{0,\omega}}{dw_1} - \frac{dz^{L_0-1}_{0,\omega}}{dw_1} \overset{L_0>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{dw_1} \qquad \frac{dz^{L_1}_{1,\omega}}{dw_1} - \frac{dz^{L_1-1}_{1,\omega}}{dw_1} \overset{L_1>0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{dw_1} - \frac{1}{\lambda w_1}$$

Together with $\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}dw_1}$, $0 \leq \ell < L_j$, these expressions imply that $\frac{dz^{\ell}_{0,\omega}}{dw_1} > 0$. $\frac{dz^{\ell}_{1,\omega}}{dw_1} > 0$ if w_1 is sufficiently high.

▶ The total number of production workers $\sum_{j=0}^{1} n^0_{j,\omega}$ decreases with wages w_1.
Follows from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}dw_1}$, with $-\frac{dq_{0,\omega}}{dw_1} - \frac{dq_{1,\omega}}{dw_1} = 0$:

$$\sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{dw_1} = -\sum_{j=0}^{1} \frac{n^0_{j,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \frac{d\bar{z}_{0,\omega}}{dw_1} < 0$$

▶ The share of CEO time $s_{j,\omega}$, local output $q_{j,\omega}$ and the number of employees production workers $n^0_{j,\omega}$ increase at the headquarters and decrease at the establishment if w_1 is sufficiently high.

1. $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}dw_1}$ and $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}dw_1}$ for $j = 0, 1$ imply, together with $\sum_{j=0}^{1} \frac{dq_{j,\omega}}{dw_1} = 0$ and $\sum_{j=0}^{1} \frac{ds_{j,\omega}}{dw_1} = 0$:

$$\frac{dq_{0,\omega}}{dw_1} = \frac{e^{\lambda z_{0,\omega}^{L_0}} e^{\lambda z_{1,\omega}^{L_1}}}{\theta_{00} e^{\lambda z_{1,\omega}^{L_1}} - \theta_{10} e^{\lambda z_{0,\omega}^{L_0}}} \left(\lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} + (1 - e^{-\lambda \bar{z_{0,\omega}}}) \right.$$
$$\left. \left(\lambda \sum_{j=0}^{1} n_{j,\omega}^0 \theta_{j0} e^{-\lambda z_{j,\omega}^{L_j}} \frac{dz_{j,\omega}^{L_j}}{dw_1} - \frac{n_{1,\omega}^0 \theta_{10} e^{-\lambda z_{1,\omega}^{L_1}}}{w_1 \left(1 - \frac{1}{f_{L_1}(\varphi_{1,\omega})}\right)} \right) \right) > 0 \quad \text{if } \frac{dz_{1,\omega}^{L_1}}{dw_1} > 0$$
$$\Rightarrow \quad \frac{dq_{1,\omega}}{dw_1} < 0$$

The condition is sufficient, but not necessary.

2. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}dw_1}$:

$$\frac{dn_{j,\omega}^0}{dw_1} = \frac{\frac{dq_{j,\omega}}{dw_1} - n_{j,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1}}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$$
$$> 0 \quad \text{for } j = 0 \text{ by } \frac{dq_{0,\omega}}{dw_1} > n_{0,\omega}^0 \lambda e^{-\lambda \bar{z}_{0,\omega}} \frac{d\bar{z}_{0,\omega}}{dw_1} \quad \text{if } \frac{dz_{1,\omega}^{L_1}}{dw_1} > 0$$
$$< 0 \quad \text{for } j = 1 \text{ by } \sum_{j=0}^{1} \frac{dn_{j,\omega}^0}{dw_1} < 0$$

3. $\frac{d^2\mathcal{L}}{d\varphi_{0,\omega}dw_1}$ yields (with a little algebra):

$$\frac{ds_{0,\omega}}{dw_1} > 0 \quad \text{by } \frac{dn_{0,\omega}^0}{dw_1} > n_{0,\omega}^0 \lambda \frac{dz_{0,\omega}^{L_0}}{dw_1} \quad \text{if } \frac{dz_{1,\omega}^{L_1}}{dw_1} > 0 \qquad \Rightarrow \quad \frac{ds_{1,\omega}}{dw_1} < 0$$

▶ The cost function $C_{0,\omega}(\tilde{q})$ and the marginal production cost $\xi_{j,\omega}$ increase with wages w_1. Follows from the envelope theorem and $\frac{d\xi_{1,\omega}}{dw_1} = \frac{d\xi_{0,\omega}}{dw_1} = \frac{d\bar{z}_{0,\omega}}{dw_1} \frac{\xi_{0,\omega}\lambda}{1-e^{-\lambda \bar{z}_{0,\omega}}}$ if $\frac{d\bar{z}_{0,\omega}}{dw_1} > 0$.

C.2.6 Proposition 5: The optimal number of layers

Both wages and helping costs are equal: $w_1 = w_0$, $\theta_{10} = \theta_{00}$.

a) To show: The average cost function of the $\{L_0/L_0\}$-organization is U-shaped in total output and reaches a minimum at $\tilde{q}^*_{L_0/L_0}$.

The firm with the $\{L_0/L_0\}$-organization chooses the same knowledge levels in both establishments by $\xi_{0,L_0/L_0} = \xi_{1,L_0/L_0}$, $\varphi_{0,L_0/L_0} = \varphi_{1,L_0/L_0}$ and $w_1 = w_0$, $\theta_{10} = \theta_{00}$. The cost function thus coincides with the cost function of a single establishment firm with $n^0_{0,L_0+1} = \sum_{j=0}^{1} n^0_{j,L_0/L_0}$. Correspondingly, Proposition 1b) applies.

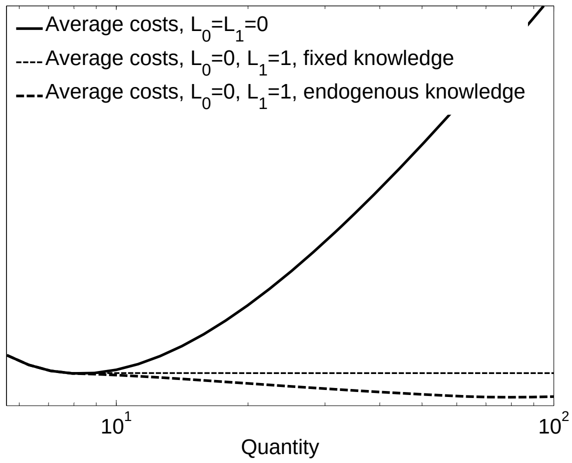
b) To show: The average cost of the $\{L_0/L_0+1\}$-organization and the $\{L_0/L_0\}$-organization are equal at $\tilde{q}^*_{L_0/L_0}$.

Follows by $q_{0,0/1} = \tilde{q}^*_{0/0}$, $q_{0,0/1} = 0$ at $\tilde{q}^*_{0/0}$.

To show: The average cost function of the $\{L_0/L_0+1\}$-organization decreases with total output $\tilde{q}$ for $\tilde{q}^*_{L_0+1/L_0+1} > \tilde{q} > \tilde{q}^*_{L_0/L_0}$.

The proof proceeds in two steps. First, we construct a ME firm organization with $\{0/1\}$ below-CEO layers and fixed knowledge levels at the minimum efficient scale of the $\{0/0\}$ organization $\tilde{q}^*_{0/0}$ and show that the organization produces

Figure C.1: Illustration: Proof of Proposition 5.



The figure illustrates part b) of the proof of the optimal number of layers of multi-establishment firms. 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 one below-CEO layer at both establishments. The dashed lines show the average cost functions of organizations with one below-CEO layer at one and two below-CEO layers at the other establishment. The light line refers to the organization with fixed knowledge levels, the bold line to the organization with endogenous knowledge levels.

$\tilde{q} \in [\tilde{q}^*_{0/0}, \tilde{q}^{MAX}]$ with constant average costs. Second, we show that the average cost function of the ME firm organization with $\{0/1\}$ below-CEO layers and endogenous knowledge levels decreases with total output for $\tilde{q} \in [\tilde{q}^*_{0/0}, \tilde{q}^*_{1/1})$, so it is lower than the average cost function of the organization with fixed knowledge levels. Figure C.1 illustrates the argument.

For simplicity and without loss of generality, we choose $L_j = 0$.

1. We construct an ME organization with $L_0 = 0$ below-CEO layers at establishment 0 and $L_1 = 1$ below-CEO layers at establishment 1 that has the same average cost as an ME organization with $L_1 = L_0 = 0$ below-CEO layers at both establishments 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 one below-CEO layer (i.e., $L = 1$). Thus, at the minimum efficient scale $\tilde{q}^*_{0/0}$,

$$\xi_{0,0/0} = \xi_{0,1} = AC_{0,1} \equiv AC^{MES}_{0,1} \tag{C.1}$$

$$\lambda z^0_{0,0/0} = \lambda z^0_{0,1} = \ln\left(\lambda \bar{z}_{0,1} + \frac{\lambda}{c}\right) + \ln\theta_{00} \equiv \lambda z^{0\,MES}_{0,1} \tag{C.2}$$

$$\lambda \bar{z}_{0,0/0} = \lambda \bar{z}_{0,1} = \lambda z^0_{0,1} + \ln\left(\lambda z^0_{0,1} + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z^0_{0,1}}\right) - \ln\theta_{00} \equiv \lambda \bar{z}^{MES}_{0,1} \tag{C.3}$$

$$\tilde{q}^*_{0,0/0} = \tilde{q}^*_{0,1} = \frac{1}{\theta_{00}} e^{\lambda z^0_{0,1}} (1 - e^{-\lambda \bar{z}_{0,1}}) \tag{C.4}$$

   Consider an ME firm with organization $L_0 = 0, L_1 = 1$, i.e., $\omega = 0/1$. Fix the knowledge levels of the firm such that

$$z^0_{0,0/1} = z^{0\,MES}_{0,1} \tag{C.5}$$

$$\bar{z}_{0,0/1} = \bar{z}^{MES}_{0,1} \tag{C.6}$$

$$1 + cz_{0,1}^{0\,MES} + \frac{c}{\lambda} = 1 + cz^0_{1,0/1} + \frac{c}{\lambda} + \frac{c}{\lambda}\theta_{11}e^{-\lambda z^0_{1,0/1}}, \quad \text{i.e. } \xi_{1,0/1} = \xi_{0,1}, \quad \text{(C.7)}$$

$$\text{and } e^{\lambda z_{0,1}^{0\,MES}} = \theta_{11}e^{\lambda(z^1_{1,0/1} - z^0_{1,0/1})}, \quad \text{i.e. } \varphi_{1,0/1} = \varphi_{0,1}, \quad \text{(C.8)}$$

$$\text{with } z^1_{1,0/1} = \frac{1}{\lambda\theta_{11}}e^{\lambda z^0_{1,0/1}} - \frac{1}{c}.$$

By construction, the average cost of the ME firm at $\tilde{q}^*_{0/0}$ are $AC_{0,0/1} = AC^{MES}_{0,1}$. The maximum producible quantity $\tilde{q}^{MAX}$ of the ME firm with organization $\omega = 0/1$ and fixed knowledge levels is given by

$$\tilde{q}^{MAX} = \frac{1}{\theta_{00}}e^{\lambda z^1_{1,0/1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}}) \quad \text{(C.9)}$$

At $\tilde{q}^{MAX}$,

$$
\begin{aligned}
\xi_{1,0/1} =& \xi_{0,0/1} = \xi_{0,1} \quad \text{by construction} && \text{(C.10)}\\
AC_{0,0/1} =& w_0 \frac{1 + cz^0_{1,0/1} + \frac{c}{\lambda} + \theta_{00}e^{-\lambda z^1_{1,0/1}}(1 + c\bar{z}^{MES}_{0,1})}{1 - e^{-\lambda \bar{z}^{MES}_{0,1}}}\\
=& \xi_{0,0/1} - w_0 \frac{\frac{c}{\lambda}\theta_{11}e^{-\lambda z^0_{1,0/1}} - \theta_{00}e^{-\lambda z^1_{1,0/1}}(1 + c\bar{z}^{MES}_{0,1})}{1 - e^{-\lambda \bar{z}^{MES}_{0,1}}}\\
\overset{\varphi_{0,0/1}=\varphi_{1,0/1}}{=}& \xi_{0,0/1} - w_0 \frac{\theta_{00}\frac{c}{\lambda}e^{-\lambda z^0_{1,0/1}}e^{-\lambda z^{0\,MES}_{0,1}}\left(e^{\lambda z^0_{0,\omega}} - \theta_{00}\left(\frac{\lambda}{c} + \lambda \bar{z}^{MES}_{0,1}\right)\right)}{1 - e^{-\lambda \bar{z}^{MES}_{0,1}}}\\
=& \xi_{0,0/1} = AC^{MES}_{0,1} \quad \text{by (C.2)} && \text{(C.11)}
\end{aligned}
$$

i.e. the ME firm produces both $\tilde{q}^*_{0/0}$ and $\tilde{q}^{MAX}$ at the same average costs.

The ME firm produces quantities $\tilde{q}$ with $\tilde{q}^{MAX} \geq \tilde{q} \geq \tilde{q}^*_{0/0}$ by allocating the share s to the establishment with one below-CEO layer and the share $1 - s$ of the production quantity to the establishment with two below-CEO layers, where

$$s = \frac{\tilde{q} - \frac{1}{\theta_{00}}e^{\lambda z^1_{1,0/1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}})}{\frac{1}{\theta_{00}}e^{\lambda z^{0\,MES}_{0,1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}}) - \frac{1}{\theta_{00}}e^{\lambda z^1_{1,0/1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}})} \quad \text{(C.12)}$$

The numerator and denominator are negative. The denominator is constant. $0 \leq s \leq 1$, because the numerator achieves its minimum at $\tilde{q} = \frac{1}{\theta_{00}}e^{\lambda z^{0\,MES}_{0,1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}})$ (so $s = 1$), and its maximum at $\tilde{q} = \frac{1}{\theta_{00}}e^{\lambda z^1_{1,0/1}}(1 - e^{-\lambda \bar{z}^{MES}_{0,1}})$ (so $s = 0$).

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 an ME firm with organization $\omega = 0/1$ and optimal knowledge levels decreases and is thus lower than the minimum average costs of the ME organization with L_0 layers at both establishments for $\tilde{q} > \tilde{q}^*_{0/0}$, because it is lower than the average cost 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 (compare the light and bold dashed line in Figure C.1) because

$$C(\tilde{q}) \leq C(\tilde{q}, \bar{z}_{0,1}^{MES}, z_{0,1}^{0\,MES}, z_{1,0/1}^{0}(z_{0,1}^{0\,MES}), z_{1,0/1}^{1}(z_{0,1}^{0\,MES})). \tag{C.13}$$

The average cost function $AC_{0,0/1}(\tilde{q})$ decreases with total output $\tilde{q}$ for $\tilde{q}^*_{1/1} > \tilde{q} > \tilde{q}^*_{0/0}$ by:

$$\begin{aligned} \frac{dAC_{0,\omega}(\tilde{q})}{d\tilde{q}} &= \frac{1}{\tilde{q}}\left(\xi_{0,\omega} - AC_{0,\omega}\right) < 0 \text{ if } \xi_{0,\omega} < AC_{0,\omega} \\ \xi_{0,0/1} = \xi_{1,0/1} < AC_{0,0/1} &\text{ if } \varphi_{0,0/1} = \varphi_{1,0/1} < w_0(1 + c\bar{z}_{0,0/1}) \end{aligned} \tag{C.14}$$

$\varphi_{j,0/1}$ is constant; $\bar{z}_{0,0/1}$ increases with $\tilde{q}$ by Proposition 3. The maximum value of $AC_{0,0/1}(\tilde{q})$ is $AC(\tilde{q}, \bar{z}_{0,1}^{MES}, z_{0,1}^{0\,MES}, z_{1,0/1}^{0}(z_{0,1}^{0\,MES}), z_{1,0/1}^{1}(z_{0,1}^{0\,MES})) = AC_{0,1}^{MES}$. At $\tilde{q}^*_{0/0}$, $\xi_{0,0/1} = AC_{0,1}$; $\xi_{0,0/1}$ decreases with $\tilde{q}$ for $\tilde{q}_{0/1}^{1/1} > \tilde{q} > \tilde{q}^*_{0/0}$. The average cost function $AC_{0,1/1}(\tilde{q})$ decreases for $\tilde{q}^*_{1/1} > \tilde{q} \geq \tilde{q}_{0/1}^{1/1}$ by Proposition 5a). □

c) To show: The average cost function of the $\{L_0+1/L_0+1\}$-organization intersects the average cost function of the $\{L_0/L_0\}$-organization at the output $\tilde{q}_{L_0/L_0}^{L_0+1/L_0+1}$ between the minimum efficient scales, i.e., $q^*_{L_0+1/L_0+1} > \tilde{q}_{L_0/L_0}^{L_0+1/L_0+1} > q^*_{L_0/L_0}$. The average cost function of the $\{L_0/L_0+1\}$-organization intersects the average cost function of the $\{L_0+1/L_0+1\}$-organization at the output $\tilde{q}_{L_0/L_0+1}^{L_0+1/L_0+1} > \tilde{q}_{L_0/L_0}^{L_0+1/L_0+1}$. We exploit the characteristics of the average cost function.

- $AC_{0,0/1} \leq AC_{0,1}^{MES}\, \forall\, \tilde{q}^*_{0/0} \leq \tilde{q} \leq \tilde{q}^{MAX}$;
- $AC_{0,0/0}$ 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 $L_0 = 0$ below-CEO layers at both establishments $AC_{0,0/0}$ intersects the decreasing average costs function of the ME firm with $L_0 = 1$ below CEO layers at both establishments $AC_{0,1/1}$ at a lower quantity than the quantity at which the decreasing average cost function of the ME firm with organization $\omega = 0/1$ $AC_{0,0/1}$ intersects the average cost function $AC_{0,1/1}$. □

Corollary 2. To show: As total output $\tilde{q}$ increases, the firm alternates between single-establishment production at location 0 and multi-establishment production with an unequal number of below-CEO layers if $w_1 = w_0$, $\theta_{10} > \theta_{00}$.

By Corollary 1, multi-establishment production with the same number of below-CEO layers at both establishments is not optimal. Multi-establishment production with a different number of below-CEO layers can be optimal for output levels between the minimum efficient scales of single-establishment firms $\tilde{q}^*_{L+1} > \tilde{q} > \tilde{q}^*_L$. Analogous to the proof of Proposition 5b), it is possible to construct an ME organization with L_0 below-CEO layers at establishment 0 and $L_0 + 1$ below-CEO layers at establishment 1 with fixed knowledge levels such that the organization has the minimum average cost of the single-establishment organization with $L = L_0 + 1$ layers for output $\tilde{q}^{MAX} \geq \tilde{q} > \tilde{q}^*_L$. The corresponding ME organization with endogenously changing knowledge levels has decreasing average costs. Multi-establishment production with a different number of below-CEO layers at the establishments thus has lower average costs than single-establishment production for a range of output levels $\tilde{q} > \tilde{q}^*_L$. □

To show: The higher the helping costs across space θ_{10} are, the lower is the range of output levels for which multi-establishment production is optimal.

Follows from Proposition 4c). □

C.2.7 Lagrangian equation and first-order conditions, firm level, local product markets

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

First-order conditions:

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

C.2.8 Proposition 6: Allocation of output/ CEO time, local product markets

To show: In optimum, $\varphi_{0,\omega} = \varphi_{1,\omega}$, and $\tau\xi_{0,\omega} = \xi_{1,\omega}$ if $q_{0,\omega} = \tilde{q}_0 + \tau(\tilde{q}_1 - q_{1,\omega})$, $\xi_{0,\omega} = \tau\xi_{1,\omega}$ if $q_{1,\omega} = \tilde{q}_1 + \tau(\tilde{q}_0 - q_{0,\omega})$ and $\xi_{0,\omega} < \tau\xi_{1,\omega} \wedge \xi_{1,\omega} < \tau\xi_{0,\omega}$ if $q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0$.

The first order conditions imply:

$\frac{\partial \mathcal{L}}{\partial s_{j,\omega}}$: If $\kappa_{j,\omega} = 0 \forall j$, $\varphi_{0,\omega} = \varphi_{1,\omega}$, see section C.2.2.

$\frac{\partial \mathcal{L}}{\partial q_{j,\omega}}$: If $\phi_{j,\omega} = 0 \forall j$,

$$\begin{aligned}\xi_{1,\omega} = \tau\xi_{0,\omega} \text{ if } q_{0,\omega} > \tilde{q}_0 \wedge q_{1,\omega} < \tilde{q}_1 \text{ by } \underline{\phi}_{0,\omega} = 0,\ \bar{\phi}_{0,\omega} = \xi_{0,\omega} = \tau^{-1}\xi_{1,\omega} \\ \xi_{0,\omega} = \tau\xi_{1,\omega} \text{ if } q_{1,\omega} > \tilde{q}_1 \wedge q_{0,\omega} < \tilde{q}_0 \text{ by } \bar{\phi}_{0,\omega} = 0,\ \underline{\phi}_{0,\omega} = \tau^{-1}\xi_{0,\omega} = \xi_{1,\omega} \\ \tau^{-1}\xi_{1,\omega} < \xi_{0,\omega} \wedge \xi_{0,\omega} < \tau\xi_{1,\omega} \text{ if } q_{0,\omega} = \tilde{q}_0 \wedge q_{1,\omega} = \tilde{q}_1 \text{ by } \underline{\phi}_{0,\omega} \neq 0 \wedge \bar{\phi}_{0,\omega} \neq 0\end{aligned}$$

If $\exists j$ s.t. $\phi_{j,\omega} > 0$, $\xi_{j,\omega} > \tau\xi_{-j,\omega}$ at $q_{j,\omega} = 0$.

C.2.9 Proposition 7: Comparative statics with respect to $\tilde{q}_j$, local product markets

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$. The second order conditions correspond to the ones in Appendix section C.2.3 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_{j,\omega}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_{j,\omega}} = \tau\frac{d\xi_{0,\omega}}{d\tilde{q}_{j,\omega}} - \frac{d\xi_{1,\omega}}{d\tilde{q}_{j,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_{0,\omega}} = 1 - \frac{dq_{0,\omega}}{d\tilde{q}_{0,\omega}} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_{0,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_{1,\omega}} = \tau - \frac{dq_{0,\omega}}{d\tilde{q}_{1,\omega}} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_{1,\omega}} = 0$$

To show: CEO knowledge $\bar{z}_{0,\omega}$ increases with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$.

1. As will be shown below, $\frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0$ and $\frac{dz^{\ell}_{j,\omega}}{d\tilde{q}} = 0$, $\ell = 0, ..., L_j$. $\frac{d^2\mathcal{L}}{dn^0_{0,\omega}d\tilde{q}_j}$ yields:

$$\frac{d\xi_{0,\omega}}{d\tilde{q}_j} = -\frac{\xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1 - e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j}$$

2. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}_k}$ $j, k = 0, 1$:

$$\frac{dn^0_{j,\omega}}{d\tilde{q}_k} = \frac{1}{1 - e^{-\lambda\bar{z}_{0,\omega}}}\left(\frac{dq_{j,\omega}}{d\tilde{q}_k} - n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_k}\right)$$

3. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}_j}$ together with $\tau\frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \frac{d\xi_{1,\omega}}{d\tilde{q}_j} = 0$ and $\mathbb{1}(j = 0)1 + \mathbb{1}(j = 1)\tau - \frac{dq_{0,\omega}}{d\tilde{q}_j} - \frac{dq_{1,\omega}}{d\tilde{q}_j} = 0$ 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: The marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$ do not vary with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$.
tba.

Case 2: $\xi_{0,\omega} = \tau\xi_{1,\omega}$. The second order conditions correspond to the ones in Appendix section C.2.3 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_{j,\omega}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_{j,\omega}} = \frac{d\xi_{0,\omega}}{d\tilde{q}_{j,\omega}} - \tau\frac{d\xi_{1,\omega}}{d\tilde{q}_{j,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_{0,\omega}} = \tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_{0,\omega}} - \frac{dq_{1,\omega}}{d\tilde{q}_{0,\omega}} = 0$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_{1,\omega}} = 1 - \tau\frac{dq_{0,\omega}}{d\tilde{q}_{1,\omega}} - \frac{dq_{1,\omega}}{d\tilde{q}_{1,\omega}} = 0$$

To show: CEO knowledge $\bar{z}_{0,\omega}$ increases with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$.
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: The marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$ do not vary with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$.

tba.

Case 3: $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$. The second order conditions correspond to the ones in Appendix section C.2.3 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_{j,\omega}} = \frac{d\xi_{0,\omega}}{d\tilde{q}_{j,\omega}} - \frac{d\bar{\phi}_{0,\omega}}{d\tilde{q}_{j,\omega}} - \tau\frac{d\underline{\phi}_{0,\omega}}{d\tilde{q}_{j,\omega}} = 0$$

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

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_{j,\omega}} = \mathbb{1}(q_{0,\omega} \geq \tilde{q}_0 \wedge q_{1,\omega} \leq \tilde{q}_1)\left(\mathbb{1}(j=0)1 + \mathbb{1}(j=1)\tau - \frac{dq_{0,\omega}}{d\tilde{q}_{j,\omega}} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_{j,\omega}}\right) = 0$$

$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_{j,\omega}} = \mathbb{1}(q_{1,\omega} \geq \tilde{q}_1 \wedge q_{0,\omega} \leq \tilde{q}_0)\left(\mathbb{1}(j=1)1 + \mathbb{1}(j=0)\tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_{j,\omega}} - \frac{dq_{1,\omega}}{d\tilde{q}_{j,\omega}}\right) = 0$$

To show: Local output $q_{j,\omega}$ varies with local demand $\tilde{q}_{j,\omega}$, but does not vary with demand at the other location $\tilde{q}_{k,\omega}$, $k \neq j$.

Follows from $\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_{j,\omega}}$ and $\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_{j,\omega}}$. For $\tilde{q}_{1,\omega}$, $\frac{dq_{1,\omega}}{d\tilde{q}_{1,\omega}} = 1 - \tau\frac{dq_{0,\omega}}{d\tilde{q}_{1,\omega}}$, so $\frac{dq_{0,\omega}}{d\tilde{q}_{1,\omega}} = \tau^2\frac{dq_{0,\omega}}{d\tilde{q}_{1,\omega}} = 0$ by $\tau^2 > 1$. For $\tilde{q}_{0,\omega}$, $\frac{dq_{0,\omega}}{d\tilde{q}_{0,\omega}} = 1 - \tau\frac{dq_{1,\omega}}{d\tilde{q}_{0,\omega}}$, so $\frac{dq_{1,\omega}}{d\tilde{q}_{0,\omega}} = \tau^2\frac{dq_{1,\omega}}{d\tilde{q}_{0,\omega}} = 0$ by $\tau^2 > 1$.

To show: CEO knowledge $\bar{z}_{0,\omega}$ increases with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$.

1. $\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\tilde{q}_k}$, $\frac{d^2\mathcal{L}}{ds_{j,\omega}d\tilde{q}_k}$, $\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\tilde{q}_k}$, $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}_k}$ and $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\tilde{q}_k}$, $j = 0, 1$ imply:

$$\frac{d\varphi_{0,\omega}}{d\tilde{q}_k} = \varphi_{0,\omega}\frac{\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} - \lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}}{(1-e^{-\lambda\bar{z}_{0,\omega}})\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}}$$

where $f_{j,\omega}(\varphi_{0,\omega}) = 1$ if $L_j = 0$, $f_{j,\omega}(\varphi_{0,\omega}) = 1 - \theta_{jj}e^{-\lambda z^0_{j,\omega}}$ if $L_j = 1$ and $f_{j,\omega}(\varphi_{0,\omega}) = 1 - \frac{e^{-\lambda z^1_{j,\omega}}}{e^{-\lambda z^0_{j,\omega}}\left(1-\theta_{jj}e^{-\lambda z^0_{j,\omega}}\right)}$ if $L_j = 2$.

2. From $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}_k}$:

$$\frac{d\xi_{j,\omega}}{d\tilde{q}_k} = \frac{1}{1-e^{-\lambda\bar{z}_{0,\omega}}}\left(\frac{d\varphi_{0,\omega}}{d\tilde{q}_k}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}} - \xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_k}\right)$$

3. Inserting $\frac{d\xi_{j,\omega}}{d\tilde{q}_k}$ and $\frac{dn^0_{j,\omega}}{d\tilde{q}_k}$ from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}_k}$ into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}_k}$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_k} = \frac{\frac{\varphi_{0,\omega}\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}} + \xi_{0,\omega}\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}}{\lambda(1+e^{-\lambda\bar{z}_{0,\omega}})\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}\sum_{j=0}^{1}\xi_{j,\omega}n^0_{j,\omega} + \varphi_{0,\omega}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}} > 0$$

To show: The marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$ do not vary with local demand $\tilde{q}_{0,\omega}$, $\tilde{q}_{1,\omega}$ if the CEO spends a sufficient share of time at the establishment with higher demand.

tba.

To show: Higher local demand $\tilde{q}_{j,\omega}$ increases (decreases) the number of production workers at the same (other) location $n^0_{j,\omega}$ ($n^0_{k,\omega}, k \neq j$).
Follows from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\tilde{q}_{j,\omega}}$ and $\frac{d^2\mathcal{L}}{d\xi_{k,\omega}d\tilde{q}_{j,\omega}}$ with $\frac{dq_{k,\omega}}{d\tilde{q}_{j,\omega}} = 0$, $\frac{dq_{j,\omega}}{d\tilde{q}_{j,\omega}} = 1$, and $0 < \frac{d\bar{z}_{0,\omega}}{d\tilde{q}_{j,\omega}} < (n^0_{j,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}})^{-1}$.

C.2.10 Proposition 8: Comparative statics with respect to θ_{10}, local product markets

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$**.** The second order conditions correspond to the ones in Appendix section C.2.4 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \tau\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}} = -\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$$

We assume that $L_0 \leq L_1$, or $L_0 > L_1$ and wages w_1 are sufficiently small.

To show: CEO knowledge $\bar{z}_{0,\omega}$ increases with the helping costs θ_{10}.

1. The two equations $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\theta_{10}}$ $j = 0, 1$ yield, 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 $\varphi_{0,\omega} = \varphi_{1,\omega}$:

$$\frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{\theta_{00}\varphi_{0,\omega} - \xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}(\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})}{(1 - e^{-\lambda\bar{z}_{0,\omega}})(\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})}$$

2. Substituting into $\frac{d^2\mathcal{L}}{dn^0_{0,\omega}d\theta_{10}}$ results in:

$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}e^{\lambda z^{L_0}_{0,\omega}}}{\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}}}$$

3. From $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$:

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

4. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\theta_{10}}$ together with $\tau\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$ and $-\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$ yields:

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

5. Combining the two expressions for $\frac{d\xi_{0,\omega}}{d\theta_{10}}$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}\theta_{00}}{\lambda\xi_{0,\omega}(1 + e^{-\lambda\bar{z}_{0,\omega}})(\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}})}$$

$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ if $\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}} > 0$. This expression holds for $L_0 \leq L_1$, see Appendix section C.2.4, and for a larger range of wages than in that section for $L_0 > L_1$.

To show: The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with the helping costs θ_{10}. Follows from $\frac{d\varphi_{1,\omega}}{d\theta_{10}} = \frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega} e^{\lambda z^{L_0}_{0,\omega}}}{\tau\theta_{00} e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10} e^{\lambda z^{L_0}_{0,\omega}}} > 0$ if $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.

To show: Local knowledge $z^{\ell}_{j,\omega}$ increases with the helping costs θ_{10}. Follows from $\frac{d\varphi_{j,\omega}}{d\theta_{10}} > 0$ and $\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega} d\theta_{10}}$, which implies:

$$\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} \overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$$\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} - \frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}} \overset{L_0>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} - \frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}} \overset{L_1>0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

To show: The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n^{\ell}_{j,\omega}$, $\forall \leq L_j$ decreases with the helping costs θ_{10}.
$\ell = 0$: Follows from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega} d\theta_{10}}$, with $-\frac{dq_{0,\omega}}{d\theta_{10}} - \tau\frac{dq_{1,\omega}}{d\theta_{10}} = 0$:

$$\sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{d\theta_{10}} = -\sum_{j=0}^{1} \frac{n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1 - e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} < 0 \quad \text{as} \quad \frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$$

$\ell > 0$: Follows from $\sum_{j=0}^{1} \frac{dn^0_{j,\omega}}{d\theta_{10}} < 0$ and $\frac{dz^{\ell}_{j,\omega}}{d\theta_{10}} > 0$.

To show: The marginal production cost $\xi_{j,\omega}$ increase with the helping costs θ_{10}. Follows from $\frac{d\xi_{1,\omega}}{d\theta_{10}} = \tau\frac{d\xi_{0,\omega}}{d\theta_{10}} = \frac{d\bar{z}_{0,\omega}}{d\theta_{10}}\frac{\xi_{0,\omega}\lambda}{1-e^{-\lambda\bar{z}_{0,\omega}}}$ 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.2.4 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} - \tau\frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\theta_{10}} = -\tau\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$$

tba.

Case 3: $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$. The second order conditions correspond to the ones in Appendix section C.2.4 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\theta_{10}} = \frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\bar{\phi}_{0,\omega}}{d\theta_{10}} - \tau\frac{d\underline{\phi}_{0,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{dq_{1,\omega}d\theta_{10}} = \frac{d\xi_{1,\omega}}{d\theta_{10}} - \tau\frac{d\bar{\phi}_{0,\omega}}{d\theta_{10}} - \frac{d\underline{\phi}_{0,\omega}}{d\theta_{10}} = 0$$
$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}} = \mathbb{1}(q_{0,\omega} \geq \tilde{q}_0 \wedge q_{1,\omega} \leq \tilde{q}_1)\left(\frac{dq_{0,\omega}}{d\theta_{10}} + \tau\frac{dq_{1,\omega}}{d\theta_{10}}\right) = 0$$
$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\theta_{10}} = \mathbb{1}(q_{1,\omega} \geq \tilde{q}_1 \wedge q_{0,\omega} \leq \tilde{q}_0)\left(\tau\frac{dq_{0,\omega}}{d\theta_{10}} + \frac{dq_{1,\omega}}{d\theta_{10}}\right) = 0$$

To show: Local output $q_{j,\omega}$ does not vary with the helping costs θ_{10}. Follows from $\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}}$ and $\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\theta_{10}}$ by $\frac{dq_{1,\omega}}{d\theta_{10}} = \tau^2\frac{dq_{1,\omega}}{d\theta_{10}}$ and $\tau^2 > 1$.

To show: CEO knowledge $\bar{z}_{0,\omega}$ increases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_1(\varphi_{0,\omega})$.

1. $\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{ds_{j,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{dz_{j,\omega}^{L_j}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$ and $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\theta_{10}}$, $j=0,1$ imply:

$$\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \varphi_{0,\omega}\frac{\frac{s_{1,\omega}}{\theta_{10}}\left(\frac{f_{1,\omega}(\varphi_{0,\omega})-1}{f_{1,\omega}(\varphi_{0,\omega})}\right) - \frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}}{\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}}$$

where $f_{j,\omega}(\varphi_{0,\omega}) = 1$ if $L_j = 0$, $f_{j,\omega}(\varphi_{0,\omega}) = 1 - \theta_{jj}e^{-\lambda z_{j,\omega}^0}$ if $L_j = 1$ and $f_{j,\omega}(\varphi_{0,\omega}) = 1 - \frac{e^{-\lambda z_{j,\omega}^1}}{e^{-\lambda z_{j,\omega}^0}\left(1-\theta_{jj}e^{-\lambda z_{j,\omega}^0}\right)}$ if $L_j = 2$.

2. From $\frac{d^2\mathcal{L}}{dn_{j,\omega}^0 d\theta_{10}}$:

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

3. Inserting $\frac{d\xi_{j,\omega}}{d\theta_{10}}$ and $\frac{dn_{j,\omega}^0}{d\theta_{10}}$ from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$ into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\theta_{10}}$ yields:

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \varphi_{0,\omega}n_{1,\omega}^0 e^{-\lambda z_{1,\omega}^{L_1}}\frac{\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})} + \frac{f_{1,\omega}(\varphi_{0,\omega})-1}{f_{1,\omega}(\varphi_{0,\omega})}}{\lambda(1+e^{-\lambda\bar{z}_{0,\omega}})\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}\sum_{j=0}^{1}\xi_{j,\omega}n_{j,\omega}^0 + \varphi_{0,\omega}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}}$$

The denominator is positive, so the sign depends on the numerator. $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ for $L_1 = 0$ because $f_{1,\omega}(\varphi_{0,\omega}) = 1$, for $L_1 = 1$ because $\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})} > 1 > \frac{1-f_{1,\omega}(\varphi_{0,\omega})}{f_{1,\omega}(\varphi_{0,\omega})}$ by $\theta_{11}e^{-\lambda z_{1,\omega}^0} \leq 0.5 \leq 1 - \theta_{11}e^{-\lambda z_{1,\omega}^0}$, and for $L_1 = 2$ if $\sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})} > \frac{1-f_{1,\omega}(\varphi_{0,\omega})}{f_{1,\omega}(\varphi_{0,\omega})}$, which holds if $s_{1,\omega} > 1 - f_{1,\omega}(\varphi_{0,\omega})$ (sufficient, not necessary).

To show: The marginal benefit of CEO time $\varphi_{j,\omega}$ decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_1(\varphi_{0,\omega})$. Follows from $f_{1,\omega}(\varphi_{0,\omega}) - 1 \leq 0$, $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ and $\frac{d\varphi_{0,\omega}}{d\theta_{10}} = \frac{d\varphi_{1,\omega}}{d\theta_{10}}$.

To show: Local knowledge $z_{j,\omega}^{\ell}$ decreases with the helping costs θ_{10} at the headquarters if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_1(\varphi_{0,\omega})$, and increases with the helping costs θ_{10} at the establishment.
$\frac{d^2\mathcal{L}}{dz_{j,\omega}^{L_j}d\theta_{10}}$ yields:

$$\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} \overset{L_0=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} \overset{L_1=0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$$\frac{dz_{0,\omega}^{L_0}}{d\theta_{10}} - \frac{dz_{0,\omega}^{L_0-1}}{d\theta_{10}} \overset{L_0>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\theta_{10}} \qquad \frac{dz_{1,\omega}^{L_1}}{d\theta_{10}} - \frac{dz_{1,\omega}^{L_1-1}}{d\theta_{10}} \overset{L_1>0}{=} \frac{1}{\varphi_{1,\omega}\lambda}\frac{d\varphi_{1,\omega}}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

$\frac{dz^{\ell}_{0,\omega}}{d\theta_{10}} < 0$ follows from $\frac{d\varphi_{0,\omega}}{d\theta_{10}} < 0$ with $\frac{d^2\mathcal{L}}{dz^{\ell}_{j,\omega}d\theta_{10}}$ for $\ell < L_0$.
$\frac{dz^{\ell}_{1,\omega}}{d\theta_{10}} > 0$ results because $\frac{\varphi_{1,\omega}}{\theta_{10}} > -\frac{d\varphi_{1,\omega}}{d\theta_{10}}$.

To show: The number of production workers $n^0_{j,\omega}$ decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_1(\varphi_{0,\omega})$.
Follows from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$ with $\frac{dq_{j,\omega}}{d\theta_{10}} = 0$ and $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$.

To show: The marginal production costs $\xi_{j,\omega}$ decreases with the helping costs θ_{10} at the headquarters and increases with the helping costs θ_{10} at the establishment.
$j = 0$: Follows from $\frac{d\varphi_{0,\omega}}{d\theta_{10}} < 0$ and $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$ for $L_1 \leq 1$. For $L_1 = 2$, a little algebra shows that the result holds if $\sum_{j=0}^{1} \frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})} > s_{1,\omega}\frac{1-f_{1,\omega}(\varphi_{0,\omega})}{f_{1,\omega}(\varphi_{0,\omega})}$, which holds by $s_{1,\omega} > s_{1,\omega}(1 - f_{1,\omega}(\varphi_{0,\omega}))$.
$j = 1$: Follows because $\varphi_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}} > \xi_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}}$, which results after substituting for $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}$ and $\frac{d\varphi_{0,\omega}}{d\theta_{10}}$ because $s_{1,\omega}\frac{1-f_{1,\omega}(\varphi_{0,\omega})}{f_{1,\omega}(\varphi_{0,\omega})} < \sum_{j=0}^{1} \frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}$.

C.3 Proposition 9: The optimal output

The profit maximization problem and the first-order conditions are given by:

$$\max_{\tilde{q}_0,\tilde{q}_1 \geq 0} \pi_i(\alpha_i) = \alpha_i^{\frac{1}{\sigma}}\left(\tilde{q}_0^{\frac{\sigma-1}{\sigma}} R_0^{\frac{1}{\sigma}} P_0^{\frac{\sigma-1}{\sigma}} + \tilde{q}_1^{\frac{\sigma-1}{\sigma}} R_1^{\frac{1}{\sigma}} P_1^{\frac{\sigma-1}{\sigma}}\right) - C(\tilde{q}_0, \tilde{q}_1)$$
$$\frac{\partial \pi_i(\alpha_i)}{\partial \tilde{q}_j} = \alpha_i^{\frac{1}{\sigma}}\frac{\sigma-1}{\sigma}\tilde{q}_j R_j^{\frac{1}{\sigma}} P_j^{\frac{\sigma-1}{\sigma}} - \xi_{j,\omega} = 0$$

$\tau\xi_{j,\omega} \neq \xi_{k,\omega}$. We define $\hat{q}_j \equiv -\tilde{q}_j$. From Proposition 8:

$$\frac{\partial^2\pi_i}{\partial\hat{q}_0\partial\theta_{10}} = \frac{\partial\xi_{0,\omega}}{\partial\theta_{10}} < 0$$
$$\frac{\partial^2\pi_i}{\partial\tilde{q}_1\partial\theta_{10}} = -\frac{\partial\xi_{1,\omega}}{\partial\theta_{10}} < 0$$

By monotone comparative statics, this implies that $\hat{q}_0$ 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 8:

$$\frac{\partial^2\pi_i}{\partial\tilde{q}_j\partial\theta_{10}} = -\frac{\partial\xi_{j,\omega}}{\partial\theta_{10}} < 0$$

From Proposition 7:

$$\frac{\partial^2\pi_i}{\partial\tilde{q}_0\partial\tilde{q}_1} = -\frac{\partial\xi_{0,\omega}}{\partial\tilde{q}_1} > 0$$

By monotone comparative statics, this implies that both $\tilde{q}_0$ and $\tilde{q}_1$ decrease with the helping costs θ_{10}. □

D Reorganization due to high-speed train routes

D.1 Robustness

Table D.1: Reduction of travel times in minutes through high speed routes

	High speed	Mean	p25	p50	p75
2000-2004	0	-1.6	-5.8	0.2	5.1
	1	-22.7	-51.5	-8.7	3.6
2004-2008	0	-1.4	-5.8	-0.2	3.1
	1	-16.8	-28.8	-9.9	-1.2

The table displays summary statistics on the reduction of travel time between 2000 and 2004 and 2004 and 2008 separately for the new high speed routes and other routes.

Table D.2: Regression results, 2000-2010 panel, drop moving establishments/ headquarters

Dep. variable	All firms				Firms with ≥ 2 establishments			
	# em.	# lay.	Mg.sh.	Mg.sh.	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.075***	0.008	0.122	−0.114	0.077***	0.001	−0.187	−0.184
	(0.011)	(0.009)	(0.251)	(0.148)	(0.012)	(0.010)	(0.266)	(0.159)
R-squared	0.892	0.863	0.833	0.872	0.893	0.867	0.834	0.872
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	88,760	88,760	88,760	88,760	79,671	79,671	79,671	79,671
# est.	12,859	12,859	12,859	12,859	11,727	11,727	11,727	11,727
Headquarters								
Firm treated	−0.010	0.031+	0.571*	−0.104	0.007	0.086***	1.502***	0.579*
	(0.015)	(0.017)	(0.276)	(0.214)	(0.021)	(0.021)	(0.343)	(0.289)
R-squared	0.948	0.887	0.932	0.897	0.952	0.894	0.936	0.904
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	20,112	20,112	20,112	20,112	10,359	10,359	10,359	10,359
# HQ	2,591	2,591	2,591	2,591	1,374	1,374	1,374	1,374
Not directly affected establishment								
Firm treated					0.006	0.005	0.446+	0.300+
					(0.013)	(0.009)	(0.256)	(0.161)
R-squared					0.926	0.889	0.872	0.886
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y
# observations					46,823	46,823	46,823	46,823
# est.					6,712	6,712	6,712	6,712

Robust standard errors in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

Table D.3: Regression results, 2000-2010 panel, bundesland-year fixed effects

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.086^{***}	0.016^{+}	0.043	−0.210	0.088^{***}	0.012	−0.004	−0.229
	(0.010)	(0.008)	(0.222)	(0.132)	(0.011)	(0.009)	(0.236)	(0.140)
R-squared	0.889	0.858	0.830	0.866	0.890	0.862	0.831	0.867
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	94,355	94,355	94,355	94,355	83,897	83,897	83,897	83,897
# est.	13,544	13,544	13,544	13,544	12,244	12,244	12,244	12,244
Headquarters								
Firm treated	-0.029^{*}	0.027^{+}	0.605^{*}	0.211	-0.039^{*}	0.052^{**}	1.213^{***}	0.804^{***}
	(0.013)	(0.014)	(0.236)	(0.178)	(0.017)	(0.017)	(0.282)	(0.282)
R-squared	0.942	0.875	0.922	0.887	0.945	0.879	0.924	0.890
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	22,972	22,972	22,972	22,972	12,487	12,487	12,487	12,487
# HQ	2,880	2,880	2,880	2,880	1,601	1,601	1,601	1,601

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

Table D.4: Regression results, 2000-2010 panel, mean travel times

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.071***	0.012	−0.095	0.114	0.062***	0.009	−0.088	0.192
	(0.011)	(0.008)	(0.237)	(0.139)	(0.012)	(0.009)	(0.254)	(0.151)
R-squared	0.890	0.859	0.832	0.868	0.891	0.864	0.833	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	94,354	94,354	94,354	94,354	83,894	83,894	83,894	83,894
# est.	13,544	13,544	13,544	13,544	12,244	12,244	12,244	12,244
Headquarters								
Firm treated	−0.015	−0.008	0.595*	0.007	−0.003	−0.001	0.843**	0.562*
	(0.013)	(0.014)	(0.249)	(0.191)	(0.018)	(0.018)	(0.321)	(0.243)
R-squared	0.945	0.882	0.926	0.892	0.950	0.889	0.931	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	22,884	22,884	22,884	22,884	12,264	12,264	12,264	12,264
# HQ	2,875	2,875	2,875	2,875	1,587	1,587	1,587	1,587
Not directly affected establishment								
Firm treated					−0.021+	0.015+	0.252	0.528***
					(0.011)	(0.008)	(0.242)	(0.147)
R-squared					0.898	0.867	0.834	0.873
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995

Robust standard errors in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

Table D.5: Regression results, 2000-2010 panel, cluster robust SE

	All firms				Firms with $\geq$ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Directly affected establishment								
Est. treated	0.074***	0.006	−0.029	−0.145	0.067***	0.000	−0.250	−0.147
	(0.015)	(0.012)	(0.353)	(0.213)	(0.020)	(0.014)	(0.345)	(0.228)
R-squared	0.890	0.859	0.832	0.868	0.891	0.864	0.833	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	94,354	94,354	94,354	94,354	83,894	83,894	83,894	83,894
# est.	13,544	13,544	13,544	13,544	12,244	12,244	12,244	12,244
Headquarters								
Firm treated	−0.020	0.018	0.537	0.063	−0.013	0.042+	0.996*	0.631
	(0.019)	(0.020)	(0.445)	(0.336)	(0.034)	(0.022)	(0.479)	(0.504)
R-squared	0.945	0.882	0.926	0.892	0.950	0.889	0.931	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
# observations	22,884	22,884	22,884	22,884	12,264	12,264	12,264	12,264
# HQ	2,875	2,875	2,875	2,875	1,587	1587	1,587	1587
Not directly affected establishment								
Firm treated					−0.030+	0.004	0.221	0.412+
					(0.017)	(0.012)	(0.281)	(0.223)
R-squared					0.898	0.867	0.834	0.873
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995

Standard errors clustered by establishment (headquarters) and establishment (headquarters) county in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

Table D.6: Regression results, 2000-2010 panel, identification as Giroud (2013)

	Firms with ≥ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)
Directly affected establishment				
Est. treated	0.070^{***}	0.007	0.011	−0.143
	(0.011)	(0.009)	(0.246)	(0.143)
R-squared	0.890	0.860	0.833	0.869
Est. FE	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y
# est.	13,526	13,526	13,526	13,526
# firms	3,003	3,003	3,003	3,003

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

Table D.7: Regression results, 2000-2010 panel, only always connected stations

	Firms with ≥ 2 establishments			
Dep. variable	# em.	# lay.	Mg.sh.	Mg.sh.
	(1)	(2)	(3)	(4)
Directly affected establishment				
Est. treated	0.077^{***}	−0.002	−0.178	-0.311^{+}
	(0.014)	(0.011)	(0.301)	(0.185)
R-squared	0.893	0.857	0.830	0.870
Est. FE	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y
# observations	63,359	63,359	63,359	63,359
# est.	9,097	9,097	9,097	9,097

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: # em.:* log number of non-managerial employees; *# lay.:* number of managerial layers of establishment/HQ; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in columns 3 and 7 and according to Blossfeld (1983) in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

D.2 Estimates at the firm level

Identification. As Figure 3 shows, only the total—direct and indirect—effect of lower helping costs on the organizational outcomes is identified. However, it is possible to back out the direct effect of lower helping costs on the number of managerial layers at the firm level. The data allow us to estimate the impact of lower helping costs on firm output and the total effect of lower helping costs on the number of layers. By using the estimate from Friedrich (2016) on the effect of output on the number of layers, we can back out the direct effect of lower helping costs on the number of layers (absent any major effects on the number of establishments).

Empirical specification. We run two sets of regressions. First, we estimate regressions similar to the establishment and headquarter level regressions in the main body of the paper:

$$y_{it} = \beta_0 + \beta_1 D_{\exists j \text{ s.t. } \theta_{j0}\downarrow,it} + \alpha_i + \alpha_{ht} + \epsilon_{it} \tag{D.1}$$

i refers to a multi-establishment firm, j to an establishment, h to the headquarter county and t indexes time. α denotes fixed effects. Unlike in the main specifications, α_{ht} refers to a headquarter county × pre/post period fixed effect, however. The number of observations, in particular for sales, is too low to estimate headquarter county × year fixed effects.

Second, we re-estimate the regressions for the number of managerial layers including the lagged number of managerial layers to make our estimates comparable to the ones in Friedrich (2016):

$$y_{it} = \beta_0 + \beta_1 D_{\exists j \text{ s.t. } \theta_{j0}\downarrow,it} + \gamma y_{it-1} + \alpha_i + \alpha_{ht} + \epsilon_{it} \tag{D.2}$$

Following Friedrich (2016), we use an Arellano-Bond estimator with three period lagged variables as instruments.

As outcome variables y_{it}, we use the sales of the firm, the number of establishments, the number of managerial layers, and the managerial shares as indicators of the organizational structure, and the number of non-managerial employees as measure of labor input.

Regression results. We first report the regression results along the lines of the estimations in the main text.

Table D.8: Regression results, 2000-2010 panel, firm level

Dep. variable	sales (1)	# em. (2)	# est. (3)	# lay. (4)	Mg.sh. (5)	Mg.sh. (6)
All firms						
Faster travel times	0.045^{+}	−0.006	−0.089	0.038^{**}	0.201	0.263^{*}
	(0.024)	(0.009)	(1.181)	(0.012)	(0.170)	(0.105)
R-squared	0.975	0.969	0.898	0.881	0.956	0.938
Firm FE	Y	Y	Y	Y	Y	Y
HQ c.-pre/post FE	Y	Y	Y	Y	Y	Y
# observations	9,529	22,975	22,975	22,975	22,975	22,975
# firms	1,827	2,874	2,874	2,874	2,874	2,874
Firms with non-missing sales						
Faster travel times	0.045^{+}	−0.024	−0.750	0.036^{*}	0.542^{*}	0.237
	(0.024)	(0.016)	(3.351)	(0.018)	(0.233)	(0.152)
R-squared	0.975	0.980	0.892	0.914	0.978	0.967
Firm FE	Y	Y	Y	Y	Y	Y
HQ c.-pre/post FE	Y	Y	Y	Y	Y	Y
# observations	9,529	9,529	9,529	9,529	9,529	9,529
# firms	1,827	1,827	1,827	1,827	1,827	1,827

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables: sales:* log sales; *# em.:* log number of non-managerial employees; *# est.:* number of establishments; *# lay.:* number of managerial layers; *Mg.sh.:* share of managerial occupations in wage sum, where managerial occupations are determined by layer in column 5 and according to Blossfeld (1983) in column 6.

The Arellano-Bond type regressions for the number of managerial layers yield:

Table D.9: Regression results, 2000-2010 panel, firm level, # managerial layers

Dep. variable	# layers (1)	# layers (2)
Faster travel times	0.030*	−0.003
	(0.012)	0.043)
Lagged # layers	0.846***	0.913***
	(0.031)	(0.040)
Sample	All firms	With sales
Firm FE	Y	Y
HQ c.-pre/post FE	Y	Y
# observations	20,983	9,437
# firms	3,023	2,069

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Direct vs. total effect. Friedrich (2016), Table VIII, contains the results of regressions of the number of layers on the lagged number of layers and log sales. The coefficient estimates of sales are between .230 and .313. Taking these estimates together with the effect of .045 of lower helping costs on log sales implies an indirect effect of lower helping costs on the number of layers of between .010 and .014. Column 2 of Table D.9 contains the effect of lower helping costs on the number of layers for the sample with non-missing sales information. As the estimated total effect is virtually zero, the implied direct effect of lower helping costs on the number of layers is between -.014 and -.010.