CESifo WORKING PAPERS

7435
2018

Original Version: December 2018
This Version: July 2019

Firm Organization with Multiple Establishments

Anna Gumpert, Henrike Steimer, Manfred Antoni

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7435*
Category 8: Trade Policy

Firm Organization with Multiple Establishments

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 a CEO's time is a resource in limited supply, shared across headquarters and establishments. Geographic frictions increase the costs of accessing the CEO. Hiring middle managers at one establishment substitutes for CEO time, which is reallocated across all establishments. Consequently, geographic frictions between the headquarters and one establishment affect the organization of all establishments of a firm. Our model is consistent with novel facts about multi-establishment firm organization that we document using administrative data from Germany. We exploit the opening of high-speed 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: D210, D220, D240.

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

Anna Gumpert
LMU Munich
Seminar for Comparative Economics
Akademiestr. 1
Germany – 80799 Munich
anna.gumpert@econ.lmu.de

Henrike Steimer
Stanford Institute for Economic Policy Research
John and Cynthia Fry Gunn Building
366 Galvez Street
USA – Stanford, CA 94305
hsteimer@stanford.edu

Manfred Antoni
Institute for Employment Research (IAB)
Regensburger Strasse 104
Germany – 90478 Nuremberg
Manfred.Antoni@iab.de

June 2019
We are deeply grateful to Lorenzo Caliendo for in-depth comments and to Monika Schnitzer for her constant support. We thank Pol Antràs, Wolfgang Dauth, and Alexandra Roulet for very helpful discussions of the paper, and Andrew Bernard, Florian Englmaier, Robert Gibbons, Martin Watzinger, Jens Wrona as well as seminar participants at Harvard, LSE, LMU Munich, MIT, Oxford, Regensburg, and the Graduate Institute, Geneva, and conference participants at the CEPR Organizational Economics Conference, NBER Organizational Economics Meeting, CESifo Area Conference on the Global Economy, Verein für Socialpolitik Annual Conference and Committee for Organizational Economics Meeting, and the European Economic Association Conference, for their comments. We thank DB Fernverkehr AG for providing the data on the travel times between cities in the German long distance railway network. The paper was partly written while Anna Gumpert was a visitor at the Yale School of Management (SOM). She thanks the SOM for its hospitality and Lorenzo Caliendo for making the visit possible. The project has benefited from financial support through the German Science Foundation (DFG) under SFB-TRR 190 and GRK 1928, and through a travel grant of the Fritz Thyssen Foundation. An early version of the paper circulated under the title "Communication and the organization of firms across space".

1 Introduction

Large firms often organize their employees in multiple establishments in different locations. Geographic frictions such as long travel times to the headquarters adversely affect establishment performance within multi-establishment firms (e.g., Giroud, 2013; Kalnins and Lafontaine, 2013). Anecdotal evidence suggests that adjusting the managerial organization may help firms mitigate the negative impacts of geographic frictions. For example, employing middle managers at regional offices instead of the headquarters proved a key ingredient for the success of Singer Sewing Machine Company in the US (Chandler, 2002, 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). Moreover, when the Canadian firm Blinds To Go set up a plant in New Jersey in 1998, moving an experienced manager on site improved its production efficiency (Menor and Mark, 2001). Nevertheless, to date, there is no theoretical work and 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 greater numbers of 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 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 other potential 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 in travel times following 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 affect not only the local establishment, but also the headquarters and other establishments of a multi-establishment firm. Local conditions thus propagate across space through firm organization.

We motivate our study by documenting three facts about multi-establishment firm organization. Our data set is ideally suited to the study of multi-establishment firms because it combines detailed information about their establishments' employees and locations. 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 particular location decreases with distance from the headquarters. Distance from headquarters also correlates negatively with the size of the establishments.

Second, the number of managerial layers of a multi-establishment firm correlates positively with the distance of its establishments from the headquarters. Quantitatively, doubling the distance is associated with the same increase in number of layers as increasing sales by 14 percent. The correlation is not driven by larger firms investing in more distant locations, or other firm characteristics. Distance correlates positively with the number of managerial layers both at the establishments and the headquarters.

Third, multi-establishment firms typically add or drop managerial layers either at the headquarters or the establishments. Only rarely do they alter the number of layers at both the headquarters and establishments simultaneously. This pattern is similar across firms with few and many establishments, and firms with close and distant establishments.

We propose a model to understand how geographic frictions affect the optimal managerial organization of firms. We model firms as knowledge hierarchies (e.g., Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). We select this framework because recent evidence suggests that distance impedes knowledge flows both within firm sites and across firms (e.g., Liu, 2010; Lychagin et al., 2016), and that the efficient transfer of intangible inputs such as managerial knowledge is an important motive for integrating multiple establishments (Atalay et al., 2014). We assume that a firm consists of a headquarters and possibly an additional establishment. The production workers at the headquarters and the establishment share a chief executive officer (CEO), who is located at the headquarters and helps workers solve the problems that arise during production. Production is a problem-solving process. Workers input labor and generate problems that must be solved using their or the CEO's knowledge in order to generate output. The firm may choose to hire a layer of local middle managers, who solve some of the problems that would otherwise need to be solved by the CEO, but entail a quasi-fixed cost for the firm.

Helping workers costs CEO time. The driving forces 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.

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

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

In the final part of our paper, we utilize the opening of high-speed train routes in Germany to study the response of firm organization to exogenous variation in geographic frictions. The train routes reduce travel time between establishments and headquarters, providing the quickest mode of travel between locations: they are faster than cars, as well as planes once 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 other potential establishments of the firm. Importantly, geographic frictions affect establishment size in the model. Size changes lead to changes in the number of layers. Travel times therefore have an indirect effect through size on the managerial organization, in addition to their direct effect. Only the total—direct and indirect—effect of lower travel times is identified.

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

Through the lens of the model, lower travel times between the headquarters and the establishment affect the managerial organization because they decrease the costs of the CEO helping the employees at the establishment. We exploit the model's implication that changes in the helping costs have a more pronounced impact in sectors with a less predictable production process to explore this channel. We find that the estimated effects are driven by establishments and headquarters in sectors with below-median predictability of the production process. This evidence supports the helping cost channel proposed by the model.

Our paper contributes to several strands of the literature. The key result of the paper is that geographic frictions are an important determinant of the managerial organization of firms. To develop this result theoretically, we build on the literature of firms as knowledge hierarchies (for an overview, see Garicano and Rossi-Hansberg, 2015). Our paper is closest to that of Antràs et al. (2008), which shows that middle managers facilitate the transmission of knowledge across countries in the context of offshoring. Our model goes beyond their theory by incorporating simultaneous production at the headquarters and the establishment of a firm. This enables us to study the effect of local shocks on the managerial organization of not only the local but also the non-local units of a firm. The broader literature on knowledge hierarchies focuses on size as a determinant of organization and shows that adding a layer of middle managers allows firms to increase efficiency as they grow (e.g., Caliendo and Rossi-Hansberg, 2012; Caliendo et al., 2015a, 2015b; Friedrich, 2016).[1] In contrast, the possibility of multi-establishment production is largely neglected, even though multi-establishment firms account for a substantial share of aggregate employment in developed economies.[2]

The key implication of the paper is that multi-establishment firm organization is interdependent across establishments. This insight is particularly relevant for recent literature documenting how multi-establishment firms propagate local shocks through their internal networks (Giroud and Mueller, 2017; Seetharam, 2018). This

[1] Similar predictions result from a monitoring hierarchy framework (e.g., Chen, 2017; Chen and Suen, 2017). Mariscal (2018) demonstrates that the impact of new information technologies on firm organization explains the decline of the US labor share using a knowledge hierarchy framework. Sforza (2017) studies the organizational responses of firms to a credit supply shock. Spanos (2018) shows that variation in firm organization explains some of the productivity differences across local markets. In the empirical literature on firm hierarchies, Rajan and Wulf (2006) document the flattening of corporate hierarchies over time, and Guadalupe and Wulf (2010) examine the impact of competition on corporate hierarchies using detailed data on 300 large publicly traded US firms.

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

body of research discusses managerial and financial constraints as potential drivers of the empirical findings. However, although CEOs are considered an important determinant of firm performance (see Bertrand, 2009, for a survey), to date managerial constraints have received very little systematic attention. Our contribution is to provide both a formal analysis and empirical evidence regarding the role of managerial constraints for multi-establishment firm organization.

Our empirical strategy builds on literature using the introduction of high-speed train routes to identify the impact of reductions in geographic frictions on firm outcomes (e.g., Bernard et al., 2017). Our empirical approach is particularly close to Charnoz et al. (2018), who to our knowledge, provide the only purely empirical study of the impact of geographic frictions on firm organization. Our theoretical model demonstrates why the impact of geographic frictions goes beyond a particular establishment, as also found by Charnoz et al. (2018), and cleanly disentangles the direct effects of geographic frictions on firm organization and the indirect effects through size. While they study business groups, our empirical analysis focuses on multi-establishment firms, i.e., the establishments are not legally independent units.

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 this assumption may apply to patents or trademarks, but not necessarily to managerial inputs.

Finally, our paper offers a novel perspective on the recent management literature. Bloom et al. (2016) document that half of the total variation in management practices between different US establishments owes to variations 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 layers of middle managers 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 final 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 the study of multi-establishment firms. The data contain information on the sales and legal form of firms, as well as location at the county level and the sector of their establishments.[3] We observe all employees per establishment subject to social security contributions on 30 June. The data include the occupation, level of education, and wages of each employee. The data cover firms in all sectors during the period 2000–2012. Each employee, establishment and firm has a unique identifier that makes it possible to follow them over time.

We assemble the data set from two sources. The universe of social security records provides the data on employees and establishments. The Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) makes these data available for research. We use the employee history, the Establishment History Panel and the extension files on entries and exits of establishments.[4] The Orbis database of Bureau van Dijk (BvD) contains the 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 must contain the firm name. The headquarters (HQ) of a firm is identified as the establishment with the same zip code or locality as the firm.[5] 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 year 2012 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 due to changes in occupational classification in that year (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 percent of the firms dropped due to this requirement have only one establishment.

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

2.2 Measures for managerial organization

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

[3]German counties are roughly comparable to counties in the US.

[4]The establishment identifier in the Establishment History Panel may change when ownership changes. The extension files render it possible to follow the establishments nonetheless.

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

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 the firm 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. Throughout the paper, we document that our findings are robust to treating the lowest level in each establishment 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. The establishments report the occupations of employees in the social security data. In multi-establishment firms, establishments may assign different occupations to similar employees. Cross-checking the results regarding the number of layers with the managerial share ensures that our results are robust to this possibility. We determine managerial occupations in two ways. On the one hand, we use the assignment of employees to layers and treat all employees above the lowest level as managerial. On the other hand, we use the Blossfeld (1983, 1987) occupational categories (see Appendix A.3 for the list of managerial occupations), which build on research from sociology and are part of the Establishment History Panel. Managers are employees in occupations that have decision-making power over the use of production factors as well as high-level officials in organizations (Blossfeld, 1983, 208). Given that German social security data do not contain the number of hours worked, and wages are censored at the social security limit, it is difficult to use other measures that are common in the literature such as the span of managerial control or wages per layer.

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 greater levels of education in the social security data. Further, we highlight how the tasks of employees systematically differ between layers in ways that plausibly reflect the 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 109,000 firms. We only observe sales for the larger firms owing to missing values in the Orbis data. The firms consist of 144,000 establishments (including

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	9.1						
Establishments	144,437	31.1						
Employees	6,356,072	34.2						
Descriptive statistics	**N**	**ME**	**Mean**	**SD**	**p25**	**p50**	**p75**	**p95**
# employees	99,545	0	42	92	13	21	39	133
	9,812	1	222	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 €.

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 as of December 2012 (Bundesagentur für Arbeit, 2016).

Nine percent of firms are multi-establishment firms. They make up a disproportionate share of establishments and employment: 31 percent of establishments belong to them, and 34 percent of employees work for them. This pattern is similar across sectors. In manufacturing—the sector with the highest number of firms—multi-establishment firms account for seven percent of firms, but 39 percent of employment. In retail and wholesale—the second largest sector—the share of multi-establishment firms is 12 percent, but their share in employment is 33 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.[6]

The statistics in the lower panel reflect the relevance of multi-establishment firms. These are substantially larger than single-establishment firms in terms of 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-establishment firms have five establishments (including headquarters). Half of them have two, and the largest five percent have 10 or more establishments (including headquarters). Multi-establishment firms are active in two sectors on average. The establishments tend to be geographically dispersed. Half of the multi-establishment firms only have establishments located within 170 km of their headquarters. At the top of the distribution, the distance between headquarters and

[6] The share of multi-establishment firms in the number of firms (employment) ranges from 4 (11) percent in construction (agriculture) to 16 (60) 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		30,117	41,725	7,025	49,915	125,253
Descriptive statistics, est.	**N**	**HQ**	**Mean**	**SD**	**p25**	**p50**	**p75**
# employees	35,080	0	32	333	2	5	16
	9,812	1	107	669	11	27	76
# managerial layers	35,080	0	1.0	0.8	0	1	2
	9,812	1	1.7	1.1	1	2	3
Managerial share	35,080	0	37	38	0	24	70
(%, layers)	9,812	1	38	32	11	32	64
Managerial share	35,080	0	8	19	0	0	5
(%, Blossfeld)	9,812	1	10	16	0	4	14

Descriptive statistics, ME firms. *# 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 and headquarters in square-kilometers; *HQ:* indicator for headquarters; *# employees:* number of full-time employees; *# managerial layers*: number of managerial layers; *Managerial share (%, layers/Blossfeld):* share of wage sum earned by employees in managerial occupations (according to layers/Blossfeld occupational categories).

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 other establishments. Headquarters are substantially larger than other establishments. The median headquarters is even larger than the median single-establishment firm. The size of establishments varies with a larger standard deviation than that for single-establishment firms. This only partly reflects that the cut-off of at least 10 employees is not binding at the establishment level. Management is concentrated in the headquarters: Headquarters have a higher number of managerial layers at all quartiles of the distribution, and a higher management share than establishments.

3 Facts

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

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

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

The table presents the coefficient estimates of a Probit model (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, one if firm i owns at least one establishment in county c, (4)-(6): log number of employees of establishment(s) in county c. *Independent variables: Log distance to HQ:* log distance between county c and HQ county of firm i in km; *Log market potential:* log of distance weighted average of the GDP of county c and surrounding counties; *Relative wages/land prices:* average wages/land prices in county c relative to wages/land prices in HQ county of firm i. We compute average wages in a county excluding firm i. Distance, market potential and relative land prices are computed using data of the German Statistical Office. The number of firms is lower than the number of ME firms due to missing values for the legal form. FE=fixed effects.

3.1 Distance to headquarters decreases location probability

Table 3 describes the geographic organization of multi-establishment firms. Columns 1 to 3 contain the results of Probit regressions that relate an indicator for whether a multi-establishment firm maintains an establishment in a county with county characteristics. Columns 4 to 6 contain OLS regressions that relate the log number of employees of the establishment(s) in a county with county characteristics.[7] The OLS regressions control for firm fixed effects to account for firm heterogeneity, hence they only include multi-establishment firms with establishments in at least two counties.

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

[7] We aggregate the number of employees to the county level if a firm has several establishments.

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 effects of land prices on location decision and size are in line with the cost of land being a fixed cost, so it is worth maintaining only larger establishments at locations with higher land prices.

Fact 1 summarizes our findings:

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

Appendix Table B.1 shows that the results are similar in the 2000–2010 panel.

3.2 Distance to headquarters increases the number of layers

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

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

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 for the fractional nature of the managerial share, we follow Papke and Wooldridge (1996) and estimate a generalized linear model using the same covariates.[8] We approximate geographic frictions with the maximum distance of an establishment to the headquarters 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 establishments in at least two counties. 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 in more distant locations.

The regression results show that both distance and area have a positive impact 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 in the number of layers as 14 percent higher sales or 19 percent more non-managerial employees. Doubling the area is associated with the same change in layers as increasing size by 30 percent according to columns 2 and 4. Accordingly, the managerial share relates positively to both the distance and the area.

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

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

Dependent variable			# managerial layers				Managerial share $\in [0, 1]$	
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.017***		0.022***		0.059***		0.040***	
	(0.004)		(0.003)		(0.009)		(0.012)	
Log area		0.025***		0.027***		0.070***		0.072***
		(0.005)		(0.004)		(0.011)		(0.013)
Log sales	0.118***	0.082***						
	(0.004)	(0.005)						
Log # non-mg. employees			0.115***	0.088***				
			(0.004)	(0.005)				
# firms	5,111	1,661	9,275	2,768	9,275	2,768	9,275	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model			Poisson				GLM	

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* (1)-(4) number of managerial layers, (5),(6) managerial share in wage sum, according to layers, (7),(8) managerial share in wage sum, according to Blossfeld occupational categories. *Independent variables: Maximum log distance to headquarters*: log of maximum distance between establishment and headquarters in km; *Log area spanned by firm*: log of minimum area covered by establishments and headquarters in square kilometers; *Log sales:* log annual sales; *Log # of non-mg. employees:* log number of employees at lowest layer. The number of observations is lower than the number of ME firms due to missing values for the legal form. FE = fixed effects, mg. = managerial.

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

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.031***	0.109***	0.086**			
	(0.007)	(0.020)	(0.033)			
Maximum log distance to HQ				0.042***	0.088***	0.069***
				(0.004)	(0.010)	(0.014)
Log # non-mg. employees	0.258***			0.169***		
	(0.010)			(0.004)		
# est./HQ	29,416	35,079	35,079	9,536	9,812	9,812
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

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

The firm-level results may disguise different responses of headquarter and establishment organization. As the descriptive statistics in Table 2 indicate, the managerial organization of the establishments does not copy the headquarters. 71 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 establishments and columns 4 to 6 refer to 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 further from the headquarters. The number of managerial layers and the managerial share at the headquarters also respond positively to the distance of the headquarters from the establishments. Quantitatively, doubling the (maximum) distance is equivalent to increasing the number of non-managerial employees by 12 percent for the establishments and 25 percent for the headquarters.

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 they span, conditional on firm characteristics. The number of managerial layers and the managerial share of both the establishments and the headquarters increase with distance.*

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

Next, we assess the robustness to alternative econometric specifications. Tables B.9 and B.10 replicate the regression results using linear models. Table B.11 shows that distance affects the managerial organization of an establishment even within firms: the number of layers and the managerial share of an establishment increase with distance in linear regressions including firm fixed effects.

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

3.3 Reorganization of headquarters or establishments

In order to complement the cross-sectional evidence on the determinants of managerial organization, we study the reorganization dynamics of firms over time. The upper panel of Table 6 displays the percentage share of multi-establishment firms that transition from a number of managerial layers in year t to a potentially different number of layers in year $t+1$. At least 80 percent of firms keep the number of layers constant across periods. If they alter the number of layers, firms add or drop one layer. These dynamics are similar to those of French and Danish firms (Caliendo et al., 2015b, Friedrich, 2016) and single-establishment firms (see Table B.18).

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

(a) # managerial layers of firm

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

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

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

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

The lower panel displays the organizational dynamics at the level of the headquarters and establishments. We count the maximum number of managerial layers at the establishments to account for the potentially different number of establishments across firms. Over time, the managerial organization at the headquarter/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, if multi-establishment firms change their organization, they typically add or drop layers at either the headquarters or the establishment(s), but not both. For example, among multi-establishment firms with two layers both at the headquarters and the establishments, nine 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 one percent of firms choose a lower or higher number of layers at both. Overall, among the firms that reorganize, 49 percent change the number of layers only at the headquarters, 42 percent change it only at the establishments, and just nine percent change it at both.[9]

Fact 3 summarizes our finding.

[9]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.

Fact 3. *Multi-establishment firms that reorganize typically add (or drop) layers either at the headquarters or at the establishments. Firms thus alternately choose organizational structures with a higher number of layers at the headquarters than the establishments and structures with an equal number of layers.*

Robustness. Appendix B.3 documents the robustness of the results. Table B.19 shows that changes in the number of layers are related to changes in firm size, consistent with the literature (Caliendo et al., 2015b, Friedrich, 2016). Table B.20 documents that the organizational dynamics are similar if the lowest-level layer in each establishment is treated as non-managerial. Table B.21 shows that the transition dynamics are comparable for longer time lags. Firms also typically add or drop layers either at the establishments or the headquarters when we consider a five-year period. Table B.22 shows that the results are similar for firms with headquarters and exactly one establishment, and for firms with headquarters and at least two establishments, hence the aggregation of the establishments in Table 6 does not drive the results. Table B.23 documents that the transition patterns are similar for firms with only proximate and firms with distant establishments.

4 Model

We use Facts 1 to 3 to inform a model in which 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 economy consists of agents and firms. The N_j agents per location each supply one unit of time to the labor market. The agents are immobile, so wages w_j may differ. Each firm i produces one product. The agents consume the products supplied on the product market.

Production. Production is a problem-solving process based on labor and knowledge (as in Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). One unit of labor 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. A higher value of λ implies that problems in the tail of the distribution occur with lower probability, so more problems can be solved with a given amount of knowledge. 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 supply knowledge, employees must learn. They spend $w_j cz$ 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 among themselves, and hence can leverage differences in knowledge. Communication is costly: an employee in location j spends θ_{kj} units of time helping an employee in location k. Helping is more costly across locations than within a 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 must 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 lowest layer. The CEO constitutes the highest managerial layer. All firms consist at least of production workers and a CEO; they may also have one or more below-CEO layers of middle managers. The knowledge levels of the employees are overlapping, so employees at layer ℓ know the knowledge of employees at layer $\ell - 1$ and more.[10] 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, who is thus a resource in limited supply for a firm.

The helping costs θ_{jk}, learning costs c, and the predictability of the production process λ are 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 examine the organization of a firm in location 0 taking output as given. Section 4.4 endogenizes output.

[10] Overlapping knowledge levels simplify the analysis, as overlaps and gaps between CEO and establishment knowledge may occur with non-overlapping knowledge in multi-establishment firms.

4.2 Single-establishment firm organization

We first determine the optimal organization of a firm in location 0 that only produces in its headquarters as a benchmark for the analysis of multi-establishment firm organization.[11] The organization consists of the number of below-CEO layers of middle managers $L \in \{0, 1, 2\}$, the number $n^{\ell}_{0,L}$ and knowledge level $z^{\ell}_{0,L}$ of employees per layer $\ell = 0, ..., L$, and the knowledge of the CEO $\bar{z}_{0,L}$.[12] The indexes $0, L$ refer to the location of the firm and the number of below-CEO layers, reflecting that these variables affect the other choices.

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

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

Output $\tilde{q}$ comprises local output $\tilde{q}_0$ and possibly the output for the other location $\tilde{q}_1$. The optimal number and knowledge levels of employees minimize costs for a given number of layers:

$$\tilde{C}_{0,L}(\tilde{q}) = \min_{\{n^{\ell}_{0,L}, z^{\ell}_{0,L}\}^{L}_{\ell=0}, \bar{z}_{0,L} \geq 0} \sum_{\ell=0}^{L} n^{\ell}_{0,L} w_0 \left(1 + c z^{\ell}_{0,L}\right) + w_0 \left(1 + c \bar{z}_{0,L}\right) \tag{2}$$

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

$$1 \geq n^{0}_{0,L} \theta_{00} e^{-\lambda z^{L}_{0,L}} \tag{4}$$

$$n^{\ell}_{0,L} \geq n^{0}_{0,L} \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} \quad \forall \ell = 1, ..., L \tag{5}$$

$$\bar{z}_{0,L} \geq z^{L}_{0,L}, \quad z^{\ell}_{0,L} \geq z^{\ell-1}_{0,L} \quad \forall \ell = 1, ..., L \tag{6}$$

The production costs consist of the personnel costs for the employees and the CEO. Constraint (3) specifies that the number of production workers and CEO knowledge must suffice to produce output $\tilde{q}$. Constraints (4) and (5) reflect that the amount of time of the CEO and the middle 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 by the helping costs, θ_{00}, and the probability that the problem is not yet solved, $e^{-\lambda z^{\ell-1}_{0,L}}$. Knowledge levels are overlapping and positive (constraint 6).

Appendix C.2.1 contains the Lagrangian equation and the first order conditions. Two multipliers from the Lagrangian equation help characterize the organization. The multiplier for constraint (3), $\xi_{0,L}$, denotes the marginal production costs. The multiplier for constraint (4), $\varphi_{0,L}$, denotes the marginal benefit of CEO time. As the shadow price, $\varphi_{0,L}$ reflects how costly the CEO time constraint is for the firm.

The firm optimally chooses CEO knowledge such that its marginal cost and its

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

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

marginal benefit are equal:

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

The marginal cost consists of the increase of CEO remuneration $w_0 c$. The marginal benefit is the lower production costs, because every unit of labor yields more output.

Given CEO knowledge, constraint (3) determines the number of production workers $n^0_{0,L}$. Constraint (4) 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 middle managers at lower layers are a recursive function of the knowledge level at the highest below-CEO layer:

$$e^{\lambda\left(z^{\ell-1}_{0,L} - z^{\ell-2}_{0,L}\right)} = \left(1 + c z^{\ell}_{0,L}\right) \frac{\lambda}{c} \quad \forall \ell = 2, ..., L, \tag{8}$$

$$e^{\lambda z^0_{0,L}} = \left(1 + c z^1_{0,L}\right) \frac{\lambda \theta_{00}}{c}. \tag{9}$$

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 highest layer. Constraint (5) 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:

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

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

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

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

a) the knowledge of the CEO $\bar{z}_{0,L}$, the number $n^{\ell}_{0,L}$ and the knowledge $z^{\ell}_{0,L}$ of the employees at all below-CEO layers $\ell \leq L$, 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 with the marginal cost function, and converges to infinity for $\tilde{q} \to 0$ and $\tilde{q} \to \infty$.

Proof. See Appendix C.2.2. □

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. The larger the output, the more problems the production workers generate and, if unsolved, communicate to higher layers. Higher output therefore increases the number of employees $n_{0,L}^{\ell}$ at all below-CEO layers. However, the amount of CEO time is fixed. The knowledge of the employees at the highest below-CEO layer $z_{0,L}^{L-1}$ increases because the CEO cannot otherwise deal with all of the problems that are communicated to him. The knowledge levels at lower layers $z_{0,L}^{\ell}$, $\ell = 0, ..., L-1$ 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, the more beneficial it would be to increase CEO time and avoid the increase in 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 costs of production increase with output.[13] On the other hand, the quasi-fixed costs of the CEO and the middle managers are spread over a larger output. For output levels below the minimum efficient scale, $\tilde{q} < \tilde{q}_L^*$, the latter effect dominates; for those 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, i.e., production possibilities. On the other hand, middle managers reduce the number of problems sent to the CEO. They thus allow decreasing the knowledge of the production workers and the marginal production costs. Consequently, adding a layer 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 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 cost function becomes flatter with the number of layers, because the marginal production costs increase less strongly with output. The firm adds a layer at the crossing $\tilde{q}_0^1$ (see also Appendix C.2.3).

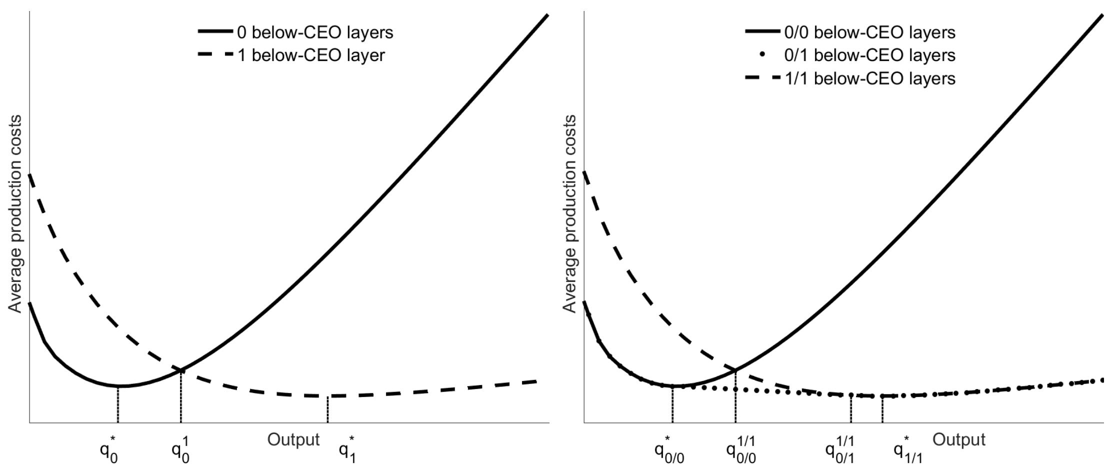
4.3 Multi-establishment firm organization

The firm may maintain an establishment at location 1 to exploit wage differences or to access the local product market. The possibility that $w_0 \geq w_1$ reflects cost-cutting motives. To capture market access motives, we assume that the firm incurs transport costs if it sells output produced in one location at the other location. We assume

[13]The marginal costs 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 function, no transport frictions

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



The figure plots the average cost functions of a SE and a ME firm for $\tau = 1$, $w_0 = w_1$, $\theta_{00} = \theta_{10}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$.
(a): The average cost function of a SE firm is U-shaped for a given number of below-CEO layers. The firm adds a layer at $\tilde{q}_0^1$.
(b): The average cost function of a ME firm with a symmetric number of below-CEO layers 0/0 or 1/1 is U-shaped. The firm adds a layer at the establishment at $\tilde{q}^*_{0/0} = \tilde{q}^*_0$ and at the headquarters at $\tilde{q}^{1/1}_{0/1} > \tilde{q}^{1/1}_{0/0} = \tilde{q}^1_0$ (or vice versa, as the 0/1 and the 1/0 organization have the same costs).

that transport costs $\tau \geq 1$ are iceberg-type.[14] We take as given the potentially different amounts of output $\{\tilde{q}_j\}_{j=0}^1$ that the firm supplies at each location.

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 layers of middle managers L_j per location. We use the term "organizational structure" and the variable ω to denote the combination of the number of below-CEO layers L_0/L_1. All other endogenous variables depend on the location and the organizational structure, so we index them by j, ω. 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 $n^\ell_{j,\omega}$ and knowledge level $z^\ell_{j,\omega}$ 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:

$$C\left(\{\tilde{q}_j\}_{j=0}^1\right) = \min_{\omega \in \Omega} \tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^1\right) \tag{10}$$

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

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

organizational structure. The costs consist of the costs per establishment and the remuneration of the CEO time that is not used in production.

$$\tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^{1}\right) = \min_{\{q_{j,\omega},s_{j,\omega}\}_{j=0}^{1},\bar{z}_{0,\omega}\geq 0} \sum_{j=0}^{1} C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0\left(1 + c\bar{z}_{0,\omega}\right) \quad (11)$$

$$\text{s.t.} \quad s_{0,\omega} + s_{1,\omega} \leq 1 \quad (12)$$

$$\mathbb{1}(q_{j,\omega} \geq \tilde{q}_j \wedge q_{k,\omega} \leq \tilde{q}_k)(q_{j,\omega} - \tilde{q}_j - \tau(\tilde{q}_k - q_{k,\omega})) \geq 0, \quad k \neq j \quad (13)$$

Equation (12) reflects the CEO's time constraint. Equation (13) states that the production quantity at location j, $q_{j,\omega}$, has to cover local output $\tilde{q}_j$ and a potential 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 $q_{k,\omega} \leq \tilde{q}_k$. In the special case that there are no transport costs, $\tau = 1$, the constraint boils down to $q_{j,\omega} + q_{k,\omega} \geq \tilde{q}_j + \tilde{q}_k \equiv \tilde{q}$, i.e., 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.

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

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

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

$$n_{j,\omega}^{\ell} \geq n_{j,\omega}^{0} \theta_{jj} e^{-\lambda z_{j,\omega}^{\ell-1}} \quad \forall \ell = 1, ..., L_j \quad (17)$$

$$\bar{z}_{0,\omega} \geq z_{j,\omega}^{L_j}, \quad z_{j,\omega}^{\ell} \geq z_{j,\omega}^{\ell-1} \quad \forall \ell = 1, ..., L_j \quad (18)$$

The constraints (15)-(18) are analogous to the constraints (3)-(6).

We solve the problem backwards. We determine the number and knowledge of the employees per layer, taking as given the firm level choices and the organizational structure. We then solve for CEO knowledge, the allocation of CEO time and the production quantities given the organizational structure, which we determine last. Appendix C.3.1 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 of CEO time) through the binding constraints (15)-(17). The formal expressions are variants of those in section 4.2, so we state them in Appendix C.3.1.

Constraint (15) determines the number of production workers. Constraint (16) 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 (17) 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 location j.

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

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

The firm optimally uses the full unit of CEO time and produces only the given output, i.e., the constraints (12) and (13) are binding. The firm can reduce production costs by reallocating CEO time as long as the marginal benefit of CEO time is not equal at the headquarters and the establishment. To some extent, the firm may also reallocate the production quantity to reduce the production costs.

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

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

The firm chooses the production quantities either exactly equal to local output or to equalize the marginal production costs adjusted by the transport costs across the headquarters and the establishment. Consequently, the marginal production costs at the headquarters and the establishment generally differ. Formally, in optimum,

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

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

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

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

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

Proof. See Appendix C.3.2. □

The firm can flexibly allocate CEO time, so it 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 production. The firm chooses between three options: it produces output locally, ships it from the other location, or does both. 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 21). If the marginal 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 22). 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 23). In the special case of no transport frictions, the firm is free to allocate production, so it reallocates quantities until the marginal costs at the headquarters and the establishment are equal. 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).

Comparative statics. To derive the optimal organizational structure ω, it is useful to understand how firm choices depend on the output $\tilde{q}_j$ and the helping costs θ_{10}. The comparative statics depend on whether the marginal costs adjusted by the transport costs are equal across the headquarters and the establishment. Parameter changes easily lead to a violation of equations (22) and (23), so these cases are unstable. The main text therefore assumes that equation (21) holds. Appendix C.3.7 contains the comparative statics for the other cases (including equation 24).

The comparative statics with respect to local output $\tilde{q}_j$ are similar to those for single-establishment firms in Proposition 1.

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

a) CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_j$. 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 $n^0_{k,\omega}$, $k \neq j$.

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

Proof. See Appendix C.3.3. □

As in Proposition 1, higher output $\tilde{q}_j$ leads to higher CEO knowledge and a higher number of production workers at location j because labor and knowledge are complementary inputs. The firm 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 higher output, the increase in 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 at both locations. Correspondingly, the marginal benefit of CEO time rises.

Higher helping costs θ_{10} at the establishment also affect the choices at the establishment and the headquarters.

Proposition 4. *Suppose the firm produces in the headquarters and the establishment. Suppose that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other, that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$, and that the helping costs across space exceed the those within a location, $\theta_{10} > \theta_{00}$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$, the knowledge of the employees at all below-CEO layers $z^{\ell}_{1,\omega}$, $\ell \leq L_1$, 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) 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 as well as the marginal benefit of CEO time $\varphi_{0,\omega}$ decrease with the helping costs θ_{10}.

These comparative statics results hold 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.3.4. □

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 must be solved at below-CEO layers, so the knowledge of the employees in the establishment increases. Therefore, the marginal production costs rise. As higher CEO knowledge yields more output per unit of labor input, the firm reduces the total number of production workers. In particular, it decreases the number of production workers at the establishment because the higher helping costs θ_{10} increase the costs of generating problems there.

The increase in helping costs for the establishment θ_{10} also affects the headquarters. The number of production workers at the headquarters decreases because local output is constant, but CEO knowledge increases. Fewer problems are generated, so the knowledge of the employees at the below-CEO layers decreases, as do the marginal production costs. The firm compensates the higher helping costs through organizational adjustments at the establishment and thus creates slack at the headquarters. Consequently, the marginal benefit of CEO time decreases.

The key implication of Propositions 3 and 4 is that the organization of the multi-establishment firm is interdependent across the headquarters and the establishment. Changes in output or helping costs at the establishment result in organizational adjustments at the establishment and the headquarters owing to the shared CEO. 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 render transparent the distinct effects of multi-establishment production and location characteristics on organizational structure, we study the organizational structure in three steps. First, we consider the special case of multi-establishment production in two locations with the same characteristics and no transport frictions, i.e., $\theta_{10} = \theta_{00}$, $w_1 = w_0$ and $\tau = 1$. Second, we add transport frictions and the possibility of geographic frictions to the helping costs, i.e., $\theta_{10} \geq \theta_{00}$, $w_1 = w_0$, $\tilde{q}_1 = \tilde{q}_0$ and $\tau > 1$. Finally, we study differences in wages and local output levels.

Special case: no transport frictions and symmetric location characteristics, $\tau = 1$, $\theta_{10} = \theta_{00}$, $w_1 = w_0$. Without transport frictions, only total output $\tilde{q}$ matters for managerial organization. The multi-establishment firm organizes production in such a way that the marginal production costs are equal at the headquarters and the establishment (Proposition 2). The marginal costs are mechanically equal if the number of below-CEO management layers is equal at the headquarters and the establishment. In this case, multi-establishment production is equivalent to single-establishment production. If the number of below-CEO management layers differ, the firm effectively produces with two distinct, albeit with the same marginal costs, production technologies. Maintaining an establishment that is different from the headquarters allows the firm to use labor and knowledge in distinct ways, and to increase output by recombining them. 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 choose the optimal combination of layers for its output by allocating CEO time and production quantities to the headquarters and the establishment. This affects the choice of optimal organizational structure.

Proposition 5. *Suppose that wages and helping costs are equal, $w_0 = w_1, \theta_{00} = \theta_{10}$, and that there are no transport frictions, $\tau = 1$. 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. 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+1/L_0-organization conincide. 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 with 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., $\tilde{q}^*_{L_0+1/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0} > \tilde{q}^*_{L_0/L_0}$. The average cost function of the L_0/L_0+1-organization intersects with the average cost function of the L_0+1/L_0+1-organization at a higher level of output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$.

As a result, the multi-establishment firm with L_0 below-CEO layers at the headquarters and the establishment adds a layer of middle managers at either the headquarters or the establishment at the output $\tilde{q}^*_{L_0/L_0}$ and at the other unit 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.3.5. □

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 functions of an organization with the same number of below-CEO layers at the headquarters and the establishment are U-shaped (Proposition 5a). They coincide with the average cost functions 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 (or 1/0-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).[15]

[15] 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.

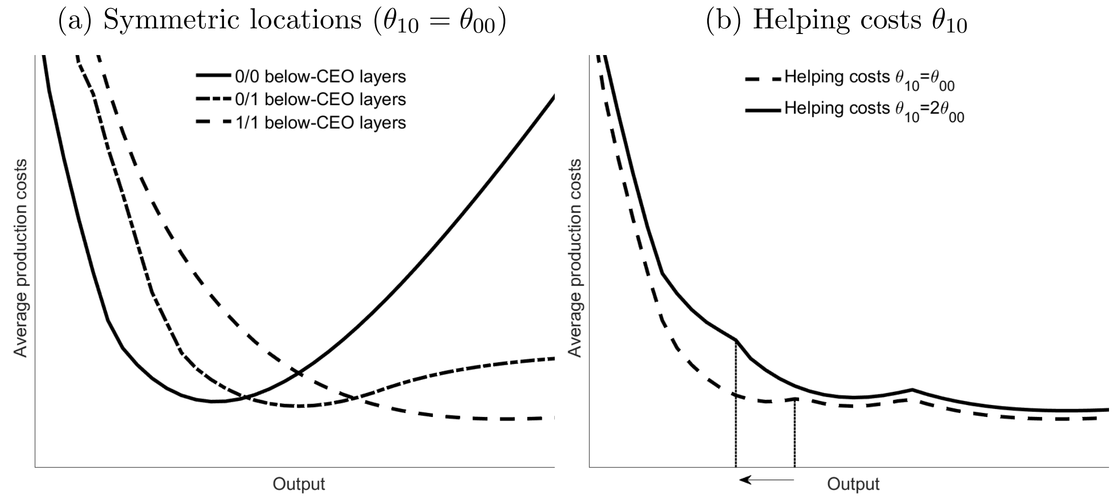
Proposition 5 is a key result of the model. It states that the multi-establishment firm successively reorganizes the headquarters and the establishment as it grows. The firm can optimally combine different numbers of below-CEO layers at the headquarters and the establishment by allocating CEO time and output and thus reduce 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 and CEO time 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 output exceeds the 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 becomes its share of production.

The layer of middle managers 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 layer thus increases efficiency both at the establishment and the headquarters. It decreases the need to add a layer of middle managers at the headquarters. As in Propositions 3 and 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 (and vice versa).

Transport and helping cost frictions, $\tau > 1$, $\theta_{10} \geq \theta_{00}$, $w_1 = w_0$, $\tilde{q}_1 = \tilde{q}_0$. Figure 2a illustrates the average production costs of different organizational structures if local output, wages and helping costs are equal at the headquarters and the establishment, but there are transport frictions. The average production costs are U-shaped, as those of a single-establishment firm. This reflects how reallocating output is efficient only under certain conditions (Proposition 2), and that the impact of higher output on firm organization is similar to its impact in a single-establishment firm (Proposition 3). As in a single-establishment firm, the minimum efficient scales increase with the number of below-CEO layers, reflecting the quasi-fixed costs of middle managers. The cost functions become flatter, because the marginal production costs increase less steeply with output due to the middle managers.

Although the transport frictions affect the shape of the average cost function of an organization with a different number of below-CEO layers at the establishment and the headquarters, they do not affect the pattern of reorganization. The multi-establishment firm does not add a layer at the headquarters and the establishment at the same size, but successively at one and the other. The additional layer of

Figure 2: Illustration of the average cost functions, transport frictions



The figure plots the average cost functions of a ME firm. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$, $w_1 = w_0$, $\tau = 1.1$, $\tilde{q}_0 = \tilde{q}_1$.
(a): At each kink, the ME firm adds a layer at one unit. The 0/1 and 1/0 organization have the same costs. (b): Higher helping costs θ_{10} decrease the output at which the firm reorganizes.

middle managers at either the headquarters or the establishment releases CEO time that can be reallocated to the establishment or the headquarters, respectively, and thus decreases the need for middle managers therein.

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 knowledge levels of employees and thus the marginal production costs at the establishment (Proposition 4). Adding a layer helps the firm to mitigate the cost increase, because it allows decreasing production worker knowledge and thus marginal costs. The higher the helping costs, the smaller is the level of output at which the firm adds a layer at the establishment, as a comparison of the solid and dashed lines demonstrates.

In addition to their effect on multi-establishment firm organization, higher helping costs affect multi-establishment production per se. Indeed, as higher helping costs increase the marginal production costs at the establishment, production there becomes less efficient relative to shipping output from the headquarters. Higher helping costs thus reduce the desirability of maintaining an establishment.

Wage differences, $w_1 \neq w_0$, $\tau > 1$, $\theta_{10} \geq \theta_{00}$ and $\tilde{q}_1 = \tilde{q}_0$. Lower wages at the establishment than the headquarters have the same effect as higher helping cost across space on the managerial organization. The firm first hires middle managers at the establishment, and then additionally at the headquarters as it grows, because it is cheaper to hire middle managers at the establishment. Higher helping costs reinforce the reorganization pattern (see Appendix Figure C.2).

Higher wages at the establishment than the headquarters render middle managers at the establishment relatively more costly. Appendix Figure C.3 illustrates the choice of organizational structure under these conditions. If only wages differ between the headquarters and the establishment, the firm adds a layer of middle managers first at the headquarters and then also at the establishment as it grows (Figure C.3a). Higher helping costs across space may outweigh the effect of higher wages at the establishment, so the firm adds a layer of middle managers only at the establishment for a range of output levels (Figure C.3b). Still, the wage difference decreases the level of output at which middle managers at both the establishment and the headquarters are optimal (Figure C.3c). Higher helping costs thus increase the number of layers of middle managers at the establishment and the headquarters.

Output differences, $\tilde{q}_1 \neq \tilde{q}_0$, $\tau > 1$, $\theta_{10} = \theta_{00}$ and $w_1 = w_0$. Appendix Figure C.4 studies output differences between the headquarters and the establishment assuming that the location characteristics are equal. We assume that headquarter output exceeds establishment output. Given that the location characteristics are the same, the analogous results hold if establishment output exceeds headquarter output. As the figure shows, hiring middle managers only at the establishment and only at the headquarters is optimal for certain output levels. As explained after Proposition 5, a multi-establishment firm with an unequal number of below-CEO layers effectively combines two different production technologies. If total output is close to the minimum efficient scale of the organization without middle managers, hiring middle managers only at the establishment—the smaller unit—is optimal, because the quasi-fixed costs of the middle managers are low, but they release CEO time and thus decrease costs at the larger unit. For higher output levels, hiring middle managers only at the headquarters—the larger unit—is optimal because the firm saves the quasi-fixed costs of middle managers at the establishment. As local output is endogenous and thus possibly depends on the helping costs, we study the interplay of higher helping costs and output in the next section.

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. We assume that each firm faces a downward sloping demand curve for its product. Firms compete monopolistically, so there is no strategic interaction between firms in their output choices.

Each firm chooses output levels $\tilde{q}_j$ to maximize profits:

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

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

Proof. See Appendix C.4.[16] □

Higher helping costs across space increase the marginal production costs at the establishment and decrease the marginal production costs at the headquarters (Proposition 4). Correspondingly, higher helping costs decrease the output at the establishment and increase the output at the headquarters. Consequently, the helping costs have both a direct effect on the managerial organization and an indirect effect through endogenous output. Thus, higher helping costs can increase the number of layers of middle managers both at the establishment and at the headquarters due to the change in the relative level of output, as Appendix Figure C.5 illustrates.

4.5 Comparison of facts and model

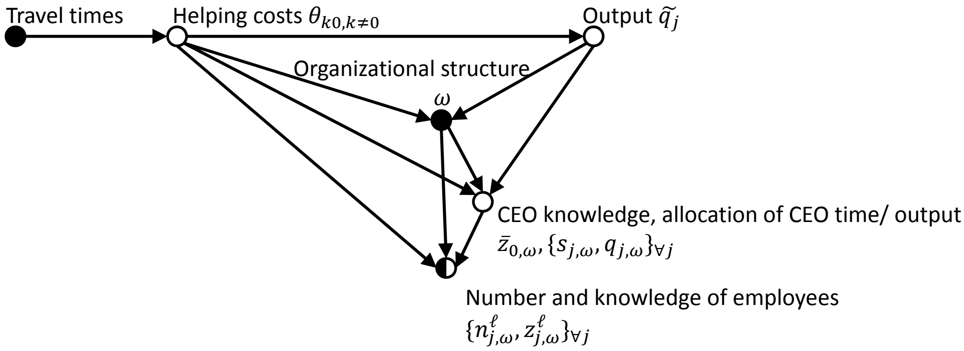
We summarize how the model relates to Facts 1 to 3. Fact 1 documents that distance to headquarters reduces the investment probability at a location and correlates negatively with establishment size. In the model, the helping costs θ_{10} reflect distance and other geographic frictions. Higher helping costs increase the marginal production costs of an establishment, and decrease its optimal output and the attractiveness of a location for an establishment.

Fact 2 documents that the number of managerial layers of a firm increases with the distance of its establishments, and that the organization of both the establishment and the headquarters responds. 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. The higher marginal costs at the establishment increase the use of middle managers there. Depending on local wages and local output, higher helping costs also increase the use of middle managers at the headquarters.

Fact 3 documents that multi-establishment firms successively add middle managers at the headquarters or the establishment as they grow. In the model, hiring middle managers at the establishment (or the headquarters) releases CEO time that is reallocated across locations. Efficiency increases throughout the firm, which reduces the need for middle managers at the headquarters (or establishment).

[16] The Appendix also includes the results for the case that $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$.

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



The graph illustrates the response of the endogenous variables to a change in 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

A key implication of the model is that geographic frictions between the establishment and the headquarters affect the organization of both the establishment and the headquarters due to the common CEO. We exploit the opening of high-speed train routes in Germany to provide evidence on this prediction.[17] The new routes make it easier to travel between the headquarters and the establishments and thus exogenously 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

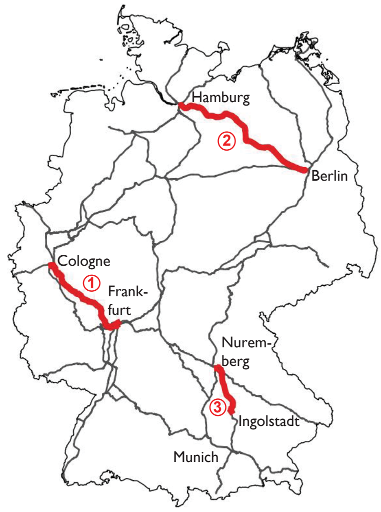
We focus on the model prediction that the helping costs θ_{10} between an establishment and the headquarters not only affect the size and organization of this establishment, but also the size and organization of the headquarters (and other possible establishments). The model disentangles how changes in helping costs affect 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.[18]

Lower travel times reduce the helping costs between an establishment k and the headquarters θ_{k0}. Lower helping costs θ_{k0} increase the optimal output $\tilde{q}_k$ at the establishment and decrease the optimal output $\tilde{q}_0$ at the headquarters. The helping

[17]Charnoz et al. (2018) and Bernard et al. (2017) also use high-speed trains for identification.

[18]To recap, optimal output is determined via profit maximization. The number and knowledge of employees are determined at the establishment level, taking as given CEO knowledge, the allocation of CEO time and production that are determined at the firm level, taking as given the organizational structure, which is determined in the final step of cost minimization.

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



The map shows the German long-distance 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).

costs thus have direct and indirect effects on firm organization. Lower helping costs directly affect the organizational structure ω, because they reduce the optimal number of layers and affect the attractiveness 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 production $q_{j,\omega}$ depend directly on θ_{k0}, but also indirectly through $\tilde{q}_j$ and ω. The choice of the number and knowledge of employees per layer $n^{\ell}_{j,\omega}, z^{\ell}_{j,\omega}$ depends directly on θ_{k0} and indirectly through $\bar{z}_{0,\omega}, s_{j,\omega}, q_{j,\omega}$ and ω.

The complexity of the relationship between the helping costs θ_{k0} and the organizational outcomes has some implications for the interpretation of the empirical estimates. The model's predictions regarding the impact of a reduction in helping costs at the establishment and headquarter level in Proposition 4 hold conditional on output and the organizational structure. These variables do not vary exogenously, but depend on the helping costs, and we do not have instruments for them. If we conditioned on output or organizational structure in an establishment-level regression, the estimation would entail a "bad control" problem (Angrist and Pischke, 2014, 214-7). Our empirical exercise therefore estimates the total—direct and indirect—effect of changes in helping costs.

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 utilize that travel times changed substantially due to the opening of three high-speed train routes during our sample period.[19] Trains on these routes exclusively transport people. Figure 4 displays a map of the new high-speed train routes and how they connect to the existing long-distance network. Deutsche Bahn AG 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 opened in May 2006 and reduced the travel time between Ingolstadt and Nuremberg 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 long-distance railway network is highly interconnected compared to other countries. For instance, Paris is the center of the French railway network, which approximately has a "star" structure. In comparison, the German network features several hubs. Reductions in travel time therefore affect more cities than merely those at the immediate ends. For example, route 1 between Cologne and Frankfurt has reduced travel times from cities in the Ruhr area to those in East and South Germany, such as Leipzig, Stuttgart, and 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 travel times as time on the train 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 ensure that temporary construction works do not affect travel times, Deutsche Bahn AG computed the travel times for three different weekdays in March, June and November. Travel times may change for several reasons, such as adjustments to time tables, construction works, or new changeover connections. To allow us to disentangle lower travel times due to the new routes and other reasons, the data contain an indicator for station pairs where more than 50 percent of passengers used one of the new 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 unobservable differences between firms connected and unconnected to the network driving the results.

A possible concern is that trains are not an attractive means of transportation for business travelers. However, this is not true of the high-speed trains. According to information from Deutsche Bahn AG for the year 2017, the share of business travelers

[19] A fourth route between Leipzig and Berlin opened in 2006. However, the travel time between these cities decreased gradually according to the data, so we cannot use this route in the estimation.

on the new routes was about double their average share.[20] This is unsurprising given that the high-speed routes render the train the fastest means of transportation between the connected cities. It is faster to travel by train than by car—it takes almost twice as long to drive from Frankfurt to Cologne than by train, for example—or even plane.[21] In addition, the high-speed trains are a flexible means of travel as regular tickets are valid on all trains that service a connection.

5.3 Empirical specification

To understand the effect of lower travel times on directly affected establishments, we estimate:[22]

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

i refers to a 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 variable of interest is an indicator variable for at least 30 minutes lower minimum travel times between the establishment and its headquarters.

To gauge the effect on the headquarters, we estimate:

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

where d denotes the headquarter county. The variable of interest is equal to one if travel times to at least one establishment decrease by at least 30 minutes. To assess the effects on non-directly affected establishments of affected firms, we estimate:

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

k refers to a non-directly affected establishment. The indicator variable is equal to one if the travel time between establishment k and the headquarters 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 the 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

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

[21] In fact, a regular plane service between Cologne Bonn Airport and Frankfurt Airport was discontinued in 2007. The carrier Lufthansa cited the new high-speed train route as the main reason for lower demand (Eurailpress.de, 2007). The number of flights between Cologne Bonn Airport and Nuremberg Airport has also dropped substantially (Deutscher Bundestag, 2007).

[22] This specification is similar to Charnoz et al. (2018).

lower travel times between the directly affected establishment and the headquarters and 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 $\mathbf{1}\{\cdot\}$ correspond to the interaction term of the "treatment" and "after" dummy variables. We implement the estimation using the `reghdfe` command by Correia (2014).

Differences in travel times may also affect other model parameters, such as local wages because employees commute longer distances (Heuermann and Schmieder, 2019). Firms may additionally 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 and establishments *in the same county and year* with constant travel times. Lower local wages or better suppliers benefit all establishments, so our estimation strategy accounts for their effect. Similarly, the regressions for the headquarters compare headquarters with travel time reductions to at least one of their establishments to headquarters *in the same county and year* without. The specification for non-directly affected establishments compares establishments that belong to firms with treated establishments to establishments in the same county and year that belong to firms without treated establishments, additionally accounting for shocks at the headquarters location.[23] Given that being treated in this set-up presupposes that firms have at least two establishments, 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 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, but this was delayed to 2004 and eventually mid-2006. Changes to establishment organization should 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 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.[24] As Appendix Table D.1

[23]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 headquarter county. However, there are too few pairs in the sample to run these regressions.

[24]One may worry that a possibly endogenous reduction in the number of changes triggers the treatment dummy. In the data, the number of changes decreases either due to the new high-

shows, virtually none of the non-high-speed-route connections exhibit lower travel times of 30 minutes or more. The threshold thus helps us ensure that the reduction is indeed driven by the exogenous new routes instead of potentially endogenous demand-driven adjustments to the time-table.

Helping cost channel. Lower travel times make it easier for headquarter managers to travel to the establishments. The model spells out a specific channel how such a reduction of within-firm geographic frictions affects the organization of multi-establishment firms: lower travel times reduce the helping costs between the headquarters and the establishment. At the same time, lower travel times may reduce other managerial frictions, such as monitoring costs (see, e.g., Giroud, 2013).

To support the helping cost channel, we document that the estimated effects are heterogeneous across sectors. We use a sector-level measure for λ, the predictability of the production process. A higher value of λ means that problems in the tail of the problem probability distribution occur with lower probability. The higher λ, the higher is the reduction in the number of problems sent to the CEO that is caused by an increase in local knowledge. Changes in helping costs should therefore have more pronounced effects on the endogenous choices in sectors with a less predictable production process, i.e., lower value of λ. Appendix D.2 contains the results of simulations corroborating this heterogeneity.

We use the measure of the predictability of the production process from Gumpert (2018). The measure is based on a survey question from the "BiBB/BAuA Employment Survey 2006" administered by the German Federal Institute for Vocational Education and Training (*Bundesinstitut für Berufsbildung*, BiBB) and the Federal Institute for Occupational Safety and Health (*Bundesanstalt für Arbeitsschutz und Arbeitsmedizin*, BAuA) (see Hall and Tiemann, 2006). The survey provides data on the employment conditions of a representative sample of 20,000 workers in Germany. The measure exploits the question of how often respondents have "to react to and solve unforeseeable problems" in their current job. It is constructed by restricting the sample to two-digit sectors with at least 25 respondents and regressing a dummy that is equal to one if participants answer "frequently," and zero if they answer "sometimes" or "never," on sector dummies. The estimated coefficients of the sector dummies are inversely related to the predictability of the production process.

We merge the measure to our data using the headquarter sector of a firm. We separately run the regressions for firms with above- and below-median predictability. While this exercise does not refute the possibility that lower travel times reduce monitoring costs within firms, it corroborates that the model mechanism of lower helping costs is at play in driving the results.

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. (1)	# lay. (2)	Mg.sh. (3)	Mg.sh. (4)	# em. (5)	# lay. (6)	Mg.sh. (7)	Mg.sh. (8)
Directly affected establishment								
Est. treated	0.061***	0.004	−0.213	−0.145	0.074***	0.001	−0.033	−0.147
	(0.018)	(0.008)	(0.226)	(0.143)	(0.019)	(0.009)	(0.241)	(0.154)
# 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
R-squared	0.802	0.874	0.912	0.868	0.803	0.878	0.911	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.015	0.019	0.688**	0.063	−0.018	0.041*	1.054**	0.631*
	(0.018)	(0.015)	(0.252)	(0.192)	(0.021)	(0.019)	(0.323)	(0.251)
# 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
R-squared	0.909	0.883	0.935	0.892	0.932	0.890	0.935	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.003	0.009	0.523*	0.412**
					(0.018)	(0.008)	(0.217)	(0.140)
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995
R-squared					0.807	0.883	0.912	0.873
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *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 in percent, where managerial occupations are determined according to layers in columns 3 and 7 and according to Blossfeld occupational categories in column 4 and 8. All variables are winsorized at the first and 99th percentiles.

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 non-directly affected establishments.

Lower travel times increase the size of the directly affected establishments. The number of non-managerial employees increases by six to seven percent. Interestingly, this increase is not accompanied by an increase in 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.

Table 8: Regression results, sample split (median predictability), 2000–2010 panel

	Below-median predictability				Above-median predictability			
Dep. variable	# em. (1)	# lay. (2)	Mg.sh. (3)	Mg.sh. (4)	# em. (5)	# lay. (6)	Mg.sh. (7)	Mg.sh. (8)
Directly affected establishment								
Est. treated	0.122***	0.017	0.350	−0.234	0.029	−0.014	−0.587*	−0.063
	(0.033)	(0.015)	(0.386)	(0.249)	(0.023)	(0.010)	(0.291)	(0.185)
# observations	44,995	44,995	44,995	44,995	45,651	45,651	45,651	45,651
# est.	7,002	7,002	7,002	7,002	6,048	6,048	6,048	6,048
R-squared	0.769	0.856	0.913	0.908	0.844	0.897	0.914	0.795
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.052+	0.065**	1.021*	0.074	0.003	−0.017	0.039	−0.118
	(0.031)	(0.024)	(0.449)	(0.362)	(0.023)	(0.022)	(0.325)	(0.232)
# observations	9,914	9,914	9,914	9,914	11,556	11,556	11,556	11,556
# HQ	1,281	1,281	1,281	1,281	1,428	1,428	1,428	1,428
R-squared	0.904	0.884	0.944	0.917	0.929	0.895	0.932	0.851
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. All variables are winsorized at the first and 99th percentiles.

Lower 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 significantly. This effect is particularly pronounced for the headquarters of firms with at least two establishments. Quantitatively, the coefficient estimates are equivalent to an increase of the managerial share by three to four percent in the average firm (six percent if defined according to the Blossfeld occupational categories). The findings are consistent with the notion that lower travel times help firms manage their growing establishments from the headquarters.

The impact of lower travel times goes beyond the headquarters. Establishments that do not themselves benefit from lower travel times, but belong to firms that do, grow at the same rate as establishments that belong to firms that do not benefit from lower travel times. However, their managerial share increases considerably, consistent with more local managerial capacity. Overall, the results strongly support the model's implication that geographic frictions between the establishment and the headquarters affect the organization of both the establishment and the headquarters as well as possible other establishments of the firm.

Helping cost channel. Table 8 documents that firms in sectors with below-median predictability of the production process drive the results in Table 7. We split the sample at median predictability. We focus on the full sample of firms and, correspondingly, on the results for establishments and headquarters. If we restrict the sample to firms with at least two establishments, the sample size decreases excessively for meaningful headquarter regressions.

As a comparison of columns 1 to 4 and columns 5 to 8 of Table 8 clearly shows, the firms in sectors with below-median predictability of the production process drive the results in Table 7. The treated establishments in those sectors grow significantly faster than untreated ones. The effect size is twice as high as in the full sample. In contrast, the treated establishments in sectors with above-median predictability grow at the same rate as untreated ones. We find organizational adjustments for the headquarters only in the low-predictability sample. The effects are again more pronounced than the corresponding effects in columns 1 to 4 of Table 7. Table 8 thus supports the notion that lower travel times affect managerial organization through the channel of lower helping costs across space.

Robustness. Appendix D documents the robustness of the results. The first set of robustness checks varies the main variables in the baseline regressions. Table D.2 documents that the results are similar if we define the treatment dummies based on the change in mean instead of minimum travel times between locations. Table D.3 shows that the results are robust if we count the managerial layers at the establishment level. One may be concerned that the regressions "overfit" the data, because they contain both establishment (or headquarter) and (headquarter) county-year fixed effects. Table D.4 replaces the (headquarter) county-year fixed effects with Bundesland-year fixed effects.[25] The Bundesland-year fixed effects reduce the number of spatial fixed effects substantially from up to 1,500 to less than 180. The estimated effects tend to be larger and slightly more significant. Table D.5 shows that the effects are robust to clustering standard errors by establishment (or headquarters) and (headquarter) county. Only for the headquarters are the effects on the managerial share now marginally significant (see also Table notes).

The second set of robustness checks alters the sample. Table D.6 replicates the regressions after dropping establishments or headquarters that move between counties. The results for the establishments are virtually unchanged; the results for the headquarters even become a little stronger. Table D.7 shows that the results are robust to restricting the sample to establishments or headquarters that are connected to the long-distance network in all years. This ensures that the high-speed routes, not (dis)connecting stations to the network, drive changes in travel time.

[25] We do not implement this check for the indirectly affected establishments, because headquarters and establishments are often in the same Bundesland.

6 Conclusion

This paper has examined the relationship between geographic frictions within firms and firm organization. We show that geographic space is an important determinant of firm organization. Using detailed and comprehensive data on German firms, we find that geographic frictions increase the use of middle managers in multi-establishment firms. Our model of managerial organization explains this finding based on the assumptions that the headquarters and establishments of a multi-establishment firm share a common CEO with limited capacity, and that geographic frictions increase the costs of accessing the CEO. The model implies that multi-establishment organization is interdependent across headquarters and establishment. We utilize the opening of new high-speed train routes as a natural experiment to provide evidence for this prediction. Consistent with the model, we find that geographic frictions between the headquarters and a single establishment not only affect the organization of that establishment, but also that of the headquarters and other establishments of the firm.

The key implication of our study is that multi-establishment organization is interdependent across establishments, hence local conditions propagate across space through firm organization. Although prior literature has discussed managerial constraints as a propagation mechanism, most analyses have prioritized the role of financial constraints. Our paper provides the first formal analysis of the role of managerial constraints for multi-establishment firm organization. Our theoretical and empirical results suggest that quantifying the relative importance of managerial and financial constraints as propagation mechanisms for local shocks would constitute a fruitful area for future research.

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, Institute for Employment Research.

Antoni, M., K. Koller, M.-C. Laible, and F. Zimmermann (2018). Orbis-ADIAB: From record linkage key to research dataset. Combining commercial company data with administrative employer-employee data. FDZ-Methodenreport 04/2018, Institute for Employment Research.

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

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

Atalay, E., A. Hortaçsu, and C. Syverson (2014). Vertical integration and input flows. *American Economic Review 104*(4), 1120–48.

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? *NBER Working Paper No. 23300*.

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

Chen, 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. *American Economic Journal: Microeconomics forthcoming*.

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. *Statistical Software Components S457874, Boston College Department of Economics*.

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, Institute for Employment Research.

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. *NBER Working Paper No. 23176*.

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 16*(6), 1929–1976.

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

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.

Liu, C. C. (2010). A spatial ecology of structural holes: scientists and communication at a biotechnology firm. In *Academy of Management Proceedings*, Volume 2010, pp. 151–155. Academy of Management Briarcliff Manor, NY 10510.

Lychagin, S., J. Pinkse, M. E. Slade, and J. Van Reenen (2016). Spillovers in space: does geography matter? *The Journal of Industrial Economics 64*(2), 295–335.

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

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

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://grlc.german-microsimulation.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. *mimeo, Stanford University*.

Sforza, A. (2017). Shocks and the organization of the firm: Who pays the bill? *mimeo, London School of Economics*.

Spanos, G. (2018). Firm Organization and Productivity Across Locations. *mimeo, University of Geneva*.

Appendix

A Data

A.1 Data sources and record linkage procedure

A.1.1 Social security data

Employee history. The Integrated Employment Biographies (*Integrierte Erwerbsbiografien, IEB*) are based on records from the German Social Security System. They contain information on all employees subject to social insurance contributions since 1975 and are updated at least annually. The data cover nearly all private sector employees in Germany, but do not include civil servants and the self-employed. The IEB contain information on birth year, gender, nationality, education, occupation, full time or part-time status and daily earnings of each employee. Jacobebbinghaus and Seth (2007) and Antoni et al. (2016) provide a detailed description of the structure of the data.[26]

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

Establishment History Panel. The Establishment History Panel (*Betriebshistorikpanel, BHP*) is a panel data set that contains information on the number of employees, sector and location of all establishments with at least one dependent employee on 30 June of each year since 1975. Following the regulations of the German Federal Employment Agency, an establishment is defined as the aggregation of all employees in a municipality that are working for the same firm in the same sector.[27] Sectors are defined based on the Classification of Economic Activities of the German Statistical Office. The location of establishments is provided at the county level. Germany is divided into 402 counties with around 200,000 inhabitants on average. German counties are roughly comparable to counties in the US. Schmucker et al. (2016) provide a detailed description of the data set.

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

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.

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

[27]That is, if a firm has several plants in a municipality, all plants in the same sector are assigned the same establishment identifier. Plants in different sectors have distinct identifiers even within the same municipality.

A.1.3 Record linkage procedure

We use a linkage table between the social security records and the Orbis database. The record linkage was performed independently of our project by the German Record Linkage Center (GRLC, see Antoni and Schnell (2017) or www.record-linkage.de for more details). The basis of the linkage was an extract of Orbis acquired by the Institute for Employment Research (IAB). This extract contained data on all German firms 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 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. Antoni et al. (2018) provide details on the process.

A.2 Sector and occupation classification

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

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

A.3 Assignment of occupations to layers

Layers. To assign occupations to layers, we build on the classification of Caliendo et al. (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. 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 layer to these occupations. Table A.2 displays our assignment of occupations to layers.

In our data, we treat the lowest level layer in each firm as non-managerial. Firms typically have consecutive layers, as the following table shows.

Table A.1: Share of firms with consecutive layers

# management layers Consecutive organization	0 Level 0	1 Level 0+1	2 Level 0+1+2	3 Level 0+1+2+3
SE firms	97%	69%	76%	100%
ME firms	96%	65%	79%	100%
Number of firms	22,068	32,573	34,130	20,586

The table displays the share of firms with consecutive layers in all firms with a given number of management layers by firm type.

Blossfeld occupational categories. The assignment from Blossfeld (1983, 1987) treats the following occupations as managerial: 751, 752, 753, 761, 762, 763.

Table A.2: 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

Continued on next page

Table A.2: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
1	31, 32, 601, 602, 603, 604, 605, 606, 607, 611, 612, 621, 622, 623, 624, 625, 626, 627, 628, 629, 633, 687, 762, 811, 812, 813, 822, 831, 841, 842, 844, 851, 852, 853, 855, 862, 863, 871, 872, 873, 874, 875, 881, 882, 883, 891, 892, 893, 922	All sub-groups of type 2 in occupation groups: 271, 273, 311, 312, 412, 414, 421, 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; 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,	Quality manager, training supervisor, management assistant, scientist, engineer, interpreter

Table A.2: 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

Evidence on wages and education by layer. Social security records.

Table A.3: Log wages and education by layer

Layer	N	Mean	p10	p25	p50	p75	p90
Level based on occupation					Log wages		
0	4,271,175	4.439	3.921	4.179	4.451	4.712	4.975
1	1,378,410	4.761	4.266	4.527	4.811	5.097	5.173
2	629,518	4.954	4.542	4.820	5.070	5.173	5.173
3	76,969	4.980	4.503	4.908	5.173	5.173	5.173
Layer in firm					Log wages		
0	4,294,004	4.440	3.922	4.180	4.452	4.714	4.977
1	1,396,933	4.763	4.270	4.530	4.814	5.099	5.173
2	605,115	4.965	4.560	4.840	5.087	5.173	5.173
3	60,020	5.004	4.551	4.988	5.173	5.173	5.173
Level based on occupation					Education		
0	4,271,175	2.002	1	2	2	2	3
1	1,378,410	2.744	2	2	2	4	4
2	629,518	3.226	2	2	4	4	4
3	76,969	3.011	2	2	3	4	4
Layer in firm					Education		
0	4,294,004	2.007	1	2	2	2	3
1	1,396,933	2.744	2	2	2	4	4
2	605,115	3.238	2	2	4	4	4
3	60,020	3.092	2	2	4	4	4

2012 cross-section. *Level based on occupation*: level of layer, see Table A.2. *Layer in firm*: layer treating the lowest level layer in each firm as non-managerial (layer 0) and assigning consecutive numbers to higher level layers. *Log wages:* log daily wages. The top percentiles at layer 1-3 are equal, 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.

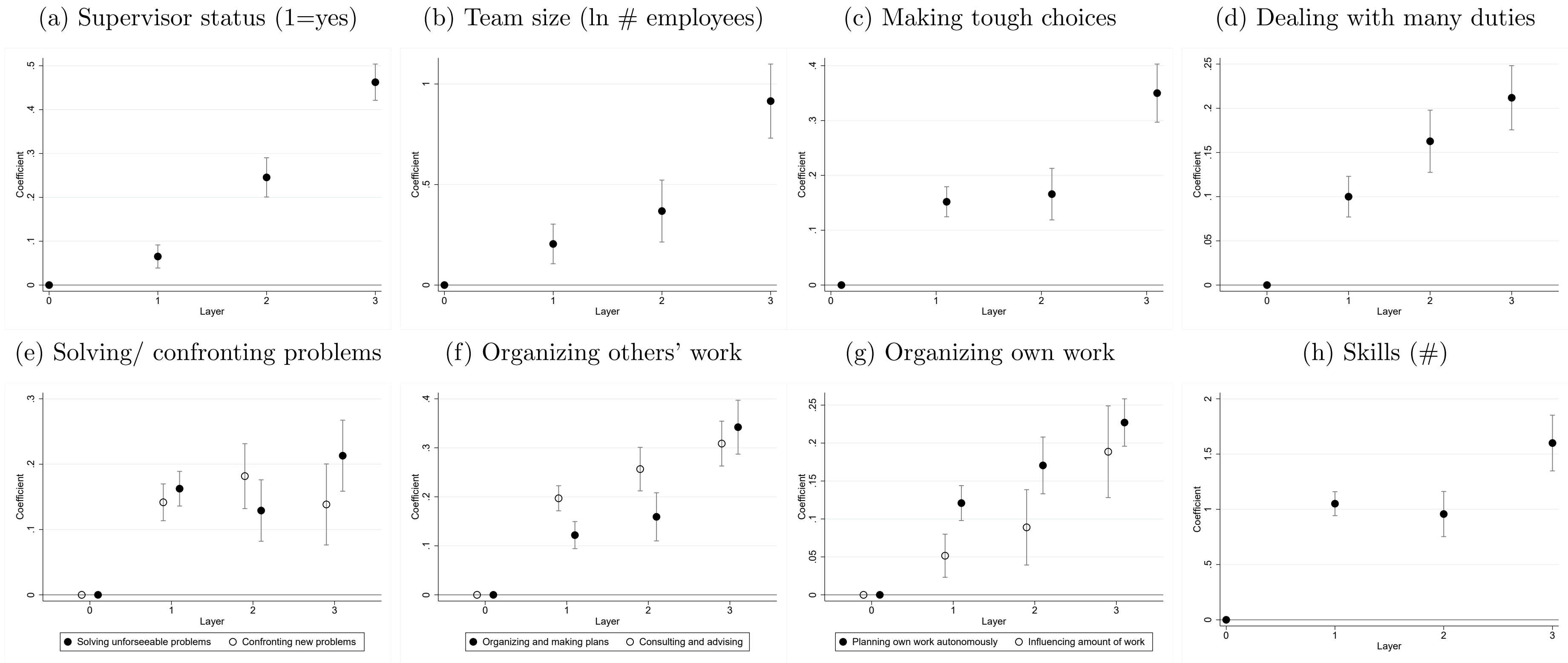
Evidence on the tasks of occupations by layer. The *2006 BiBB/BAuA Survey of the Working Population* administered by the German Federal Institute for Vocational Education and Training (*Bundesinistitut für Berufsbildung, BiBB*) and the Federal Institute for Occupational Safety and Health (*Bundesanstalt 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. Table A.4 presents the regression results.

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



The figure plots the estimated coefficients of the layer dummies in equation (A.1). In figures (c)-(g), 1=often. See notes of Table A.4 for the survey questions.

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

Table A.4: Regression results: tasks by layer, 2006 BiBB/BAuA survey

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

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

B Facts

B.1 Distance to headquarters decreases location probability

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

Dependent variable	Location probability			Log # est. employees		
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	-0.315^{***}	-0.314^{***}	-0.379^{***}	-0.109^{***}	-0.107^{***}	-0.146^{***}
	(0.020)	(0.021)	(0.016)	(0.022)	(0.021)	(0.019)
Log market potential	0.744^{***}	0.752^{***}		0.521^{***}	0.526^{***}	
	(0.025)	(0.029)		(0.061)	(0.052)	
Relative wages	-1.121^{***}	-0.859^{***}		-0.385^{*}	-0.363^{*}	
	(0.057)	(0.059)		(0.172)	(0.159)	
Relative land prices		-0.016^{*}			0.011	
		(0.006)			(0.007)	
# observations	24,393,507	11,832,114	24,393,507	171,146	86,084	171,146
# firms	10,323	8,478	10,323	8,547	6,982	8,547
HQ sector FE	Y	Y	Y	N	N	N
HQ county FE	Y	Y	Y	N	N	N
Legal form FE	Y	Y	Y	N	N	N
County FE	N	N	Y	N	N	Y
Year FE	Y	Y	Y	N	N	N
Firm-year FE	N	N	N	Y	Y	Y
Model		Probit			OLS	

The table presents the coefficient estimates of a probit model (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. Land prices are only available from 2005. Columns 4-6 only includes firms establishments in at least two counties. FE = fixed effects.

B.2 Distance to headquarters increases number of layers

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

Dependent variable	# managerial. layers				Mg. share $\in [0,1]$ Layers	
	(1)	(2)	(3)	(4)	(5)	(6)
Maximum log distance to HQ	0.018*** (0.004)		0.019*** (0.004)		0.051*** (0.008)	
Log area		0.025*** (0.005)		0.026*** (0.004)		0.068*** (0.010)
Log sales	0.127*** (0.004)	0.093*** (0.006)				
Log # non-mg. employees			0.140*** (0.004)	0.113*** (0.006)		
# firms	5,111	1,661	9,275	2,768	9,275	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM	

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

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

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

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 and 2, robust in columns 3 and 4). + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* (1),(3) number of managerial layers, (2),(4) managerial share in wage sum, both defined treating the lowest-level layer in each establishment/the HQ as non-managerial. *Independent variables:* see Table 5.

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

Dependent variable	# managerial layers		Mg. share ∈ [0, 1]	
			Layers	Blossfeld
	(1)	(2)	(3)	(4)
Quartile 2	0.018	0.024	−0.012	−0.009
	(0.018)	(0.015)	(0.035)	(0.045)
Quartile 3	0.064***	0.071***	0.174***	0.120**
	(0.018)	(0.015)	(0.035)	(0.045)
Quartile 4	0.082***	0.112***	0.324***	0.240***
	(0.018)	(0.015)	(0.037)	(0.047)
Log sales	0.117***			
	(0.004)			
Log # non-mg. employees		0.113***		
		(0.004)		
# firms	5,111	9,275	9,275	9,275
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y
Model	Poisson		GLM	

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

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

Unit	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 quartile 2	0.086***	0.260***	0.107			
	(0.020)	(0.062)	(0.091)			
Log distance quartile 3	0.126***	0.443***	0.310*			
	(0.024)	(0.079)	(0.138)			
Log distance quartile 4	0.125***	0.414***	0.413**			
	(0.026)	(0.074)	(0.128)			
Max. log distance quartile 2				0.085***	0.086*	0.084
				(0.018)	(0.039)	(0.058)
Max. log distance quartile 3				0.132***	0.230***	0.180**
				(0.018)	(0.040)	(0.057)
Max. log distance quartile 4				0.185***	0.427***	0.372***
				(0.018)	(0.041)	(0.056)
Log # non-mg. employees	0.258***			0.169***		
	(0.010)			(0.004)		
# est./HQ	29,416	35,079	35,079	9,536	9,812	9,812
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* see Table 5. *Independent variables: Log distance quartile 2-4*: dummies for quartiles of log of distance between establishment and headquarters in km; *Max. log distance quartile 2-4*: dummies for quartiles of log of maximum distance between establishment and headquarters in km; *Log # of non-mg. employees:* log number of employees at lowest layer.

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

Dependent variable	# managerial layers			
	(1)	(2)	(3)	(4)
Maximum log	0.014***		0.021***	
distance to HQ	(0.004)		(0.003)	
Log area		0.019***		0.027***
		(0.004)		(0.004)
Log sales	0.465***	0.461***		
	(0.030)	(0.041)		
Log sales, squared	−0.016***	−0.017***		
	(0.001)	(0.002)		
Log # non-mg.			0.091***	0.109***
employees			(0.012)	(0.017)
Log # non-mg.			0.003*	−0.002
employees, squared			(0.001)	(0.002)
# firms	5,111	1,661	9,275	2,768
HQ sector FE	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y

The table presents the estimated coefficients of Poisson regressions. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* number of managerial layers. *Independent variables:* see Table 4.

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

Unit	Est.	HQ
Dependent variable	# mg. layers	
	(1)	(2)
Log distance	0.031***	
to HQ	(0.007)	
Maximum log		0.043***
distance to HQ		(0.004)
Log # non-mg.	0.342***	0.230***
employees	(0.024)	(0.014)
Log # non-mg.	−0.015***	−0.009***
employees, squared	(0.003)	(0.002)
# est./HQ	29,416	9,536
Sector FE	Y	Y
County FE	Y	Y

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

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

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.017***		0.022***		0.059***		0.040***	
	(0.004)		(0.003)		(0.009)		(0.012)	
Log area		0.025***		0.027***		0.069***		0.072***
		(0.005)		(0.004)		(0.011)		(0.013)
Log sales	0.118***	0.083***						
	(0.004)	(0.005)						
Log # non-mg. employees			0.116***	0.088***				
			(0.004)	(0.005)				
# establishments	−0.000	−0.000	−0.000*	−0.000	0.000	0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# firms	5,111	1,661	9,275	2,768	9,275	2,768	9,275	2,768
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent variable:* see Table 4. *Independent variables: # establishments:* number of establishments (excluding HQ), others: see Table 4.

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

Dependent variable	# managerial layers				Mg. share (%)			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.028***		0.040***		1.124***		0.300***	
	(0.008)		(0.006)		(0.168)		(0.089)	
Log area		0.050***		0.054***		1.325***		0.555***
		(0.011)		(0.009)		(0.236)		(0.116)
Log sales	0.254***	0.192***						
	(0.007)	(0.013)						
Log # non-mg. employees			0.244***	0.200***				
			(0.007)	(0.013)				
# firms	5,066	1,529	9,253	2,673	9,253	2,673	9,253	2,673
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table 4.

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

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share (%)		# layers	Mg. share (%)	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.030***	2.036***	0.554*			
	(0.006)	(0.386)	(0.221)			
Maximum log distance to HQ				0.070***	1.715***	0.606***
				(0.007)	(0.191)	(0.123)
Log # non-mg. employees	0.279***			0.310***		
	(0.008)			(0.007)		
# est./HQ	29,396	35,061	35,061	9,514	9,790	9,790
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y

Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table 5.

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

Dependent variable	# layers	Mg. share (%)	
		Layers	Blossfeld
	(1)	(2)	(3)
Log distance to HQ	0.012+	0.432*	0.013
	(0.006)	(0.202)	(0.159)
Log # non-mg. employees	0.253***		
	(0.009)		
# est./HQ	23,698	28,828	28,828
Sector FE	Y	Y	Y
County FE	Y	Y	Y
Firm FE	Y	Y	Y

Standard errors clustered by firm in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table 5.

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

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.035***		0.029***		0.066***		0.035***	
	(0.005)		(0.004)		(0.010)		(0.014)	
Log area		0.032***		0.038***		0.058***		0.047*
		(0.005)		(0.005)		(0.013)		(0.018)
Log sales	0.144***	0.115***						
	(0.004)	(0.007)						
Log # non-mg. employees			0.193***	0.152***				
			(0.005)	(0.007)				
# firm-years	22,417	7,383	40,609	12,257	40,609	12,257	40,609	12,257
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficients. Constant included. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with at least two establishments (plus headquarters). *Dependent and independent variables:* see Table 4.

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

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	−0.016***	0.061***	−0.002			
	(0.004)	(0.010)	(0.018)			
Maximum log distance to HQ				0.051***	0.077***	0.059***
				(0.004)	(0.010)	(0.016)
Log # non-mg. employees	0.351***			0.254***		
	(0.015)			(0.005)		
# est./HQ-years	315,661	331,391	331,391	77,131	77,715	77,715
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

Standard errors clustered by firm in parentheses. *** $p < 0.001$. *Dependent and independent variables:* see Table 5.

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

Dependent variable	# managerial layers				Mg. share $\in [0,1]$			
					Layers		Blossfeld	
Horizontal	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.021** (0.007)		0.025*** (0.005)		0.059*** (0.013)		0.061*** (0.017)	
Log area		0.019* (0.008)		0.027*** (0.007)		0.071*** (0.020)		0.060** (0.023)
Log sales	0.128*** (0.006)	0.105*** (0.011)						
Log # non-mg. employees			0.120*** (0.006)	0.100*** (0.012)				
# firms	2,548	716	4,705	1,249	4,705	1,249	4,705	1,249
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

Vertical	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.012* (0.005)		0.021*** (0.004)		0.059*** (0.012)		0.021 (0.016)	
Log area		0.022*** (0.006)		0.021*** (0.006)		0.057*** (0.016)		0.075*** (0.017)
Log sales	0.114*** (0.005)	0.070*** (0.006)						
Log # non-mg. employees			0.111*** (0.005)	0.076*** (0.006)				
# firms	2,563	945	4,570	1,519	4,570	1,519	4,570	1,519
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The table presents the coefficient estimates. Constant included. Robust standard errors in parentheses. *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table 4. *Horizontal* restricts the sample to firms with all establishments in the same sector as the HQ. *Vertical* restricts the sample to firms with at least one establishment in a different sector than the HQ.

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

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share ∈ $[0, 1]$		# layers	Mg. share ∈ $[0, 1]$	
		Layers	Blossfeld		Layers	Blossfeld
Horizontal	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.040***	0.134***	0.122***			
	(0.010)	(0.026)	(0.028)			
Maximum log distance to HQ				0.045***	0.087***	0.070**
				(0.007)	(0.015)	(0.022)
Log # non-mg. employees	0.267***			0.164***		
	(0.011)			(0.007)		
# est./HQ	17,248	20,454	20,454	4,844	4,983	4,983
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

Vertical	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.021*	0.071**	0.041			
	(0.009)	(0.027)	(0.040)			
Maximum log distance to HQ				0.040***	0.099***	0.079**
				(0.006)	(0.013)	(0.018)
Log # non-mg. employees	0.252***			0.174***		
	(0.013)			(0.006)		
# est./HQ	12,168	14,625	14,625	4,692	4,829	4,829
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

The table presents the coefficient estimates. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table 5. *Horizontal* restricts the sample to establishments in the HQ sector and HQ of firms with all establishments in the HQ sector, respectively. *Vertical* restricts the sample to establishments in a different sector than the HQ and HQ of firms with at least one establishment in a different sector than the HQ, respectively.

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

Dependent variable	# managerial layers				Mg. share $\in [0, 1]$			
					Layers		Blossfeld	
GmbH	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.070***		0.017^{+}		0.044*		0.031	
	(0.013)		(0.010)		(0.022)		(0.026)	
Log area		0.015		0.049***		0.133***		0.090*
		(0.016)		(0.013)		(0.033)		(0.036)
Log sales	0.149***	0.088**						
	(0.013)	(0.018)						
Log # non-mg. employees			0.177***	0.125***				
			(0.011)	(0.024)				
# firms	724	215	1,493	452	1,493	452	1,493	452
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

GmbH & Co. KG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.014**		0.022***		0.061***		0.051***	
	(0.005)		(0.004)		(0.010)		(0.014)	
Log area		0.033***		0.029***		0.069***		0.064***
		(0.006)		(0.005)		(0.013)		(0.016)
Log sales	0.139***	0.120***						
	(0.005)	(0.007)						
Log # non-mg. employees			0.144***	0.119***				
			(0.005)	(0.008)				
# firms	3,979	1,212	7,214	2,018	7,214	2,018	7,214	2,018
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

AG	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	−0.007		−0.003		0.003		−0.066	
	(0.009)		(0.011)		(0.033)		(0.070)	
Log area		−0.007		0.008		−0.004		0.014
		(0.011)		(0.011)		(0.049)		(0.056)
Log sales	0.023**	0.020^{+}						
	(0.008)	(0.011)						
Log # non-mg. employees			0.050***	0.039***				
			(0.009)	(0.010)				
# firms	397	228	549	291	549	291	549	291
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Model	Poisson				GLM			

The tables present the coefficient estimates separately for firms with the legal form *GmbH*, *GmbH & Co. KG* and *AG*. Constant included. Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Even columns include only ME firms with establishments in at least two counties. *Dependent and independent variables:* see Table 4. 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 B.17: Regression results, mg. organization of establishments, by legal form, 2012 cross-section

Unit	Establishment			Headquarters		
Dependent variable	# layers	Mg. share $\in [0,1]$		# layers	Mg. share $\in [0,1]$	
		Layers	Blossfeld		Layers	Blossfeld
GmbH	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.044** (0.017)	−0.002 (0.052)	0.139** (0.050)			
Maximum log distance to HQ				0.058*** (0.012)	0.100*** (0.028)	0.044 (0.046)
Log # non-mg. employees	0.330*** (0.025)			0.200*** (0.013)		
# est./HQ	4,746	5,304	5,304	1,464	1,493	1,493
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM	Poisson	GLM		

GmbH & Co. KG	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.036*** (0.008)	0.113*** (0.025)	0.102* (0.046)			
Maximum log distance to HQ				0.043*** (0.005)	0.091*** (0.012)	0.092*** (0.017)
Log # non-mg. employees	0.262*** (0.011)			0.182*** (0.006)		
# est./HQ	17,905	20,754	20,754	7,005	7,214	7,214
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y
Model	Poisson	GLM		Poisson	GLM	

AG	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	0.007 (0.010)	0.102** (0.038)	0.037 (0.035)			
Maximum log distance to HQ				−0.005 (0.014)	−0.032 (0.042)	−0.069 (0.065)
Log # non-mg. employees	0.219*** (0.021)			0.060*** (0.012)		
# est./HQ	3,832	5,764	5,764	532	549	549
Sector FE	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y

The tables present the coefficient estimates separately for establishments/HQ of firms with the legal form *GmbH*, *GmbH & Co. KG* and *AG*. Constant included. Standard errors in parentheses (clustered by firm in columns 1 to 3, robust in columns 4 to 6). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent and independent variables:* see Table 4. 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.

B.3 Reorganization of headquarters or establishments

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

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

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

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

(a) firm

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

(b) headquarters/establishments

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

Panel (a) displays the average log number of employees of firms that transition from a number of managerial layers in year t (given in the rows) to a potentially different number of layers in year $t+1$ (given in the columns). Panel (b) displays the average log number of employees of firms that transition from a managerial organization in year t (given in the rows) to a potentially different managerial organization in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of layers of the headquarters. The figure behind the slash denotes the maximum number of layers at the establishments. Firms with a higher number of layers at the establishments than at the HQ are dropped for readability. The stars denote whether average size of firms that change their organization is significantly different from the average size of those that do not (marked in bold). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. .. denotes cells with fewer than 50 observations. Empty cells contain fewer than .5% of firms. Fewer than .5% of firms exit. Unreported results with sales as outcome variable are similar.

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

(a) # managerial layers of firm

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

(b) # managerial layers at headquarters/establishment

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

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

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

(a) # managerial layers of firm

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

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

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

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

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

(a) ME firms with headquarters and one establishment

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

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

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

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

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

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

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

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

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

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

C Model

C.1 Set-up

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.2 Single-establishment firm organization

C.2.1 Lagrangian equation and first order conditions

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

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

$$\frac{\partial \mathcal{L}}{\partial \bar{z}_{0,L}} = w_0 c - \xi_{0,L} n^0_{0,L} \lambda e^{-\lambda \bar{z}_{0,L}} - \bar{\eta}^L_{0,L} = 0$$

$$\frac{\partial \mathcal{L}}{\partial z^L_{0,L}} \begin{cases} \overset{L=0}{=} n^0_{0,0}\left(w_0 c - \varphi_{0,0}\theta_{00}\lambda e^{-\lambda z^0_{0,0}}\right) + \bar{\eta}^1_{0,0} - \bar{\eta}^0_{0,0} = 0 \\ \overset{L\geq 0}{=} n^0_{0,L}\left(w_0 c \theta_{00} e^{-\lambda z^{L-1}_{0,L}} - \varphi_{0,L}\theta_{00}\lambda e^{-\lambda z^L_{0,L}}\right) + \bar{\eta}^{L+1}_{0,L} - \bar{\eta}^L_{0,L} = 0 \end{cases}$$

$$\frac{\partial \mathcal{L}}{\partial z^{\ell}_{0,L}} = n^0_{0,L} w_0 \left(c\theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} - \lambda\theta_{00} e^{-\lambda z^{\ell}_{0,L}}(1 + c z^{\ell+1}_{0,L})\right) - \bar{\eta}^{\ell}_{0,L} + \bar{\eta}^{\ell+1}_{0,L} = 0$$

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

$$\frac{\partial \mathcal{L}}{\partial z^0_{0,L}} \overset{L\geq 0}{=} n^0_{0,L} w_0 \left(c - \lambda\theta_{00} e^{-\lambda z^0_{0,L}}(1 + c z^1_{0,L})\right) + \bar{\eta}^1_{0,L} - \bar{\eta}^0_{0,L} = 0$$

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

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

$$\frac{\partial \mathcal{L}}{\partial \varphi_{0,L}} = n^0_{0,L}\theta_{00} e^{-\lambda z^L_{0,L}} - 1 = 0$$

C.2.2 Proposition 1: Comparative statics

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

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

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

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

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 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 $\frac{d^2\mathcal{L}}{dz^\ell_{0,L}d\tilde{q}}$, $\ell < L$:
$$\frac{d\varphi_{0,0}}{d\tilde{q}} = \varphi_{0,0}\lambda\frac{dz^0_{0,0}}{d\tilde{q}} \equiv \varphi_{0,0}\lambda f_0(\varphi_{0,0})\frac{dz^L_{0,L}}{d\tilde{q}}$$
$$\frac{d\varphi_{0,1}}{d\tilde{q}} = \varphi_{0,1}\lambda\frac{dz^1_{0,1}}{d\tilde{q}}(1-\theta_{00}e^{-\lambda z^0_{0,1}}) \equiv \varphi_{0,1}\lambda f_1(\varphi_{0,1})\frac{dz^L_{0,L}}{d\tilde{q}}$$
$$\frac{d\varphi_{0,2}}{d\tilde{q}} = \varphi_{0,2}\lambda\frac{dz^2_{0,2}}{d\tilde{q}}\left(1-\frac{\frac{dz^1_{0,2}}{d\tilde{q}}}{\frac{dz^2_{0,2}}{d\tilde{q}}}\right) \equiv \varphi_{0,2}\lambda f_2(\varphi_{0,2})\frac{dz^L_{0,L}}{d\tilde{q}}$$
$f_L(\varphi_{0,L}) > 0$: $f_0(\varphi_{0,0}) = 1 > 0$; $f_1(\varphi_{0,1}) = 1-\theta_{00}e^{-\lambda z^0_{0,1}} > 0$; $f_2(\varphi_{0,2}) = 1 - \frac{dz^1_{0,2}}{d\tilde{q}}\Big/\frac{dz^2_{0,2}}{d\tilde{q}} > 0$ if $\frac{dz^2_{0,2}}{d\tilde{q}} > \frac{dz^1_{0,2}}{d\tilde{q}}$. This is the case if $e^{-\lambda z^1_{0,2}} < e^{-\lambda z^0_{0,2}}(1-\theta_{00}e^{-\lambda z^0_{0,2}})$ $(*)$. $(*)$ holds for $z^2_{0,2} = z^1_{0,2}$, as we can rewrite $(*)$ as $\theta_{00}e^{-\lambda z^0_{0,2}} < 1-\theta_{00}e^{-\lambda z^0_{0,2}}$ for $z^2_{0,2} = z^1_{0,2}$. $(*)$ holds for $z^2_{0,2} > z^1_{0,2}$, as both sides of $(*)$ decrease in $z^0_{0,2}$ and the left-hand side of $(*)$ decreases at a faster rate than the right hand-side of $(*)$ at $z^2_{0,2} = z^1_{0,2}$ by $-\lambda e^{-\lambda z^0_{0,2}} < -\lambda e^{-\lambda z^0_{0,2}}(1-2\theta_{00}e^{-\lambda z^0_{0,2}})$.
5. Substituting into $\frac{d^2\mathcal{L}}{d\bar{z}_{0,L}d\tilde{q}}$ yields:

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

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

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

Knowledge of employees:

1. $\ell = L$: $\frac{dz_{0,L}^{L}}{d\tilde{q}} = \frac{1}{\lambda n_{0,L}^{0}} \frac{dn_{0,L}^{0}}{d\tilde{q}} > 0$ by $\frac{dn_{0,L}^{0}}{d\tilde{q}} > 0$.
2. $\ell = 0$, $L > 0$: $\frac{dz_{0,L}^{0}}{d\tilde{q}} = \theta_{00} e^{-\lambda z_{0,L}^{0}} \frac{dz_{0,L}^{1}}{d\tilde{q}} > 0$ by $\frac{dz_{0,L}^{1}}{d\tilde{q}} > 0$.
3. $\ell = 1$, $L = 2$: $\frac{dz_{0,2}^{1}}{d\tilde{q}} = \frac{dz_{0,2}^{2}}{d\tilde{q}} \frac{e^{-\lambda z_{0,2}^{1}}}{e^{-\lambda z_{0,2}^{0}}(1 - \theta_{00} e^{-\lambda z_{0,2}^{0}})} > 0$ by $\frac{dz_{0,2}^{2}}{d\tilde{q}} = \frac{dz_{0,L}^{L}}{d\tilde{q}} > 0$. □

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

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

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

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

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

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

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

C.2.3 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 (8) and (9) define the optimal knowledge levels recursively:

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

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

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

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:

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

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

$$
\begin{aligned}
\text{MES}: &\quad \frac{dz_{0,L}^L}{d\bar{z}_{0,L}} = \frac{1}{\lambda\bar{z}_{0,L} + \frac{\lambda}{c}}\frac{1}{f_L(\varphi_{0,L})} && > 0 \\
\text{FOC, } L = 0: &\quad \frac{dz_{0,0}^0}{d\bar{z}_{0,0}} = \frac{\lambda z_{0,0}^0 + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z_{0,0}^0}}{\lambda z_{0,0}^0 + \frac{\lambda}{c} + 1 + f_0(\varphi_{0,0})} && > 0 \\
\text{FOC, } L = 1: &\quad \frac{dz_{0,1}^1}{d\bar{z}_{0,1}} = \frac{\lambda z_{0,1}^0 + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z_{0,1}^0} + \theta_{00}e^{-\lambda z_{0,1}^1}}{\lambda z_{0,1}^0 + \frac{\lambda}{c} + 1 + \theta_{00}e^{-\lambda z_{0,1}^0} + \theta_{00}e^{-\lambda z_{0,1}^0}f_1(\varphi_{0,1})} && > 0 \\
\text{FOC, } L = 2: &\quad \frac{dz_{0,2}^2}{d\bar{z}_{0,2}} = \frac{\lambda z_{0,2}^0 + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z_{0,2}^{\ell}} + \theta_{00}e^{-\lambda z_{0,2}^2}}{\lambda z_{0,2}^0 + \frac{\lambda}{c} + 1 + \theta_{00}\sum_{\ell=0}^{1} e^{-\lambda z_{0,2}^{\ell}} + \theta_{00}e^{-\lambda z_{0,2}^1}f_2(\varphi_{0,2})} && > 0
\end{aligned}
$$

where $f_L(\varphi_{0,L})$ is defined in section C.2.2.

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

C.3 Multi-establishment firm organization

C.3.1 Lagrangian equation and first order conditions

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

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

First-order conditions:

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

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

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

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

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

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.3.2 Proposition 2: Allocation of output and CEO time

To show: In optimum, $\varphi_{0,\omega} = \varphi_{1,\omega}$.
$\frac{\partial \mathcal{L}}{\partial s_{j,\omega}}$: If $\kappa_{j,\omega} = 0 \forall j$, i.e. if the CEO spends positive time on both establishments,

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

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

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

To show: In optimum, $\tau\xi_{0,\omega} = \xi_{1,\omega}$ 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 $q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0$.
$\frac{\partial \mathcal{L}}{\partial q_{j,\omega}}$: If $\phi_{j,\omega} = 0 \forall j$, i.e. if there is positive production at both establishments,

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

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

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

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

To show: If $\tau = 1$, in optimum, $\xi_{0,\omega} = \xi_{1,\omega}$.
$\frac{\partial \mathcal{L}}{\partial q_{j,\omega}}$ for $\tau = 1$: If $\phi_{j,\omega} = 0 \, \forall j$, i.e. if there is positive production at both establishments,

$$
\begin{aligned}
\phi_{0,\omega} &= \frac{\partial C}{\partial q_{0,\omega}} - \mathbb{1}(q_{0,\omega} \geq \tilde{q}_0 \wedge q_{1,\omega} \leq \tilde{q}_1)\bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \geq \tilde{q}_1 \wedge q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{0,\omega} \\
&= \frac{\partial C}{\partial q_{1,\omega}} - \mathbb{1}(q_{0,\omega} \geq \tilde{q}_0 \wedge q_{1,\omega} \leq \tilde{q}_1)\bar{\phi}_{0,\omega} - \mathbb{1}(q_{1,\omega} \geq \tilde{q}_1 \wedge q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{0,\omega} = \phi_{1,\omega} = 0 \\
&\quad \text{for } q_{0,\omega}, q_{1,\omega} > 0 \\
\Rightarrow &\quad \frac{\partial C}{\partial q_{0,\omega}} = \xi_{0,\omega} = \xi_{1,\omega} = \frac{\partial C}{\partial q_{1,\omega}}. \qquad \square
\end{aligned}
$$

C.3.3 Proposition 3: Comparative statics with respect to $\tilde{q}_j$

We substitute $\frac{d\mathcal{L}}{dz^{\ell}_{k,\omega}}$, $\ell \leq L_k$, into $\frac{d^2\mathcal{L}}{dn^0_{k,\omega}d\tilde{q}_j}$. We assume that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$ and $q_{j,\omega}$, $s_{j,\omega} > 0\,\forall j$. The second order conditions are, with $k, j \in \{0, 1\}$:

$$
\begin{aligned}
\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_j} &= \frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \frac{d\bar{\phi}_{0,\omega}}{d\tilde{q}_j} - \tau\frac{d\underline{\phi}_{0,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_j} &= \frac{d\xi_{1,\omega}}{d\tilde{q}_j} - \tau\frac{d\bar{\phi}_{0,\omega}}{d\tilde{q}_j} - \frac{d\underline{\phi}_{0,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{ds_{0,\omega}d\tilde{q}_j} - \frac{d^2\mathcal{L}}{ds_{1,\omega}d\tilde{q}_j} &= \frac{d\varphi_{0,\omega}}{d\tilde{q}_j} - \frac{d\varphi_{1,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{d\bar{z}_{0,\omega}d\tilde{q}_j} &= -\sum_{k=0}^{1}\frac{d\xi_{k,\omega}}{d\tilde{q}_j}n^0_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}} - \sum_{k=0}^{1}\xi_{k,\omega}\frac{dn^0_{k,\omega}}{d\tilde{q}_j}\lambda e^{-\lambda\bar{z}_{0,\omega}} \\
&\quad + \sum_{k=0}^{1}\xi_{k,\omega}n^0_{k,\omega}\lambda^2 e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{d\kappa_{0,\omega}d\tilde{q}_j} &= \frac{ds_{0,\omega}}{d\tilde{q}_j} + \frac{ds_{1,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_j} &= \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} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_j}\right) = 0 \\
\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_j} &= \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} - \frac{dq_{1,\omega}}{d\tilde{q}_j}\right) = 0 \\
\frac{d^2\mathcal{L}}{dz^{L_k}_{k,\omega}d\tilde{q}_j} &\begin{cases} \overset{L_k=0}{=} -\frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}\lambda e^{-\lambda z^0_{k,\omega}} + \varphi_{k,\omega}\theta_{k0}\lambda^2 e^{-\lambda z^0_{k,\omega}}\frac{dz^0_{k,\omega}}{d\tilde{q}_j} = 0 \\ \overset{L_k>0}{=} -w_k c\theta_{kk}\lambda e^{-\lambda z^{L_k-1}_{k,\omega}}\frac{dz^{L_k-1}_{k,\omega}}{d\tilde{q}_j} - \frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}\lambda e^{-\lambda z^{L_k}_{k,\omega}} + \varphi_{k,\omega}\theta_{k0}\lambda^2 e^{-\lambda z^{L_k}_{k,\omega}}\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} = 0 \end{cases} \\
\frac{d^2\mathcal{L}}{dz^{\ell}_{k,\omega}d\tilde{q}_j} &= -\lambda c e^{-\lambda z^{\ell-1}_{k,\omega}}\frac{dz^{\ell-1}_{k,\omega}}{d\tilde{q}_j} + \lambda^2 e^{-\lambda z^{\ell}_{k,\omega}}\frac{dz^{\ell}_{k,\omega}}{d\tilde{q}_j}(1 + cz^{\ell+1}_{k,\omega}) - \lambda e^{-\lambda z^{\ell}_{k,\omega}}c\frac{dz^{\ell+1}_{k,\omega}}{d\tilde{q}_j} = 0 \\
&\quad \text{for } 0 < \ell < L_k,\ L_k > 1 \\
\frac{d^2\mathcal{L}}{dz^0_{k,\omega}d\tilde{q}_j} &\overset{L_k>0}{=} \lambda^2\theta_{kk}e^{-\lambda z^0_{k,\omega}}\frac{dz^0_{k,\omega}}{d\tilde{q}_j}(1 + cz^1_{k,\omega}) - \lambda\theta_{kk}e^{-\lambda z^0_{k,\omega}}c\frac{dz^1_{k,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{dn^0_{k,\omega}d\tilde{q}_j} &= -\frac{d\xi_{k,\omega}}{d\tilde{q}_j}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - \xi_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} + \frac{d\varphi_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} = 0 \\
\frac{d^2\mathcal{L}}{d\xi_{k,\omega}d\tilde{q}_j} &= \frac{dq_{k,\omega}}{d\tilde{q}_j} - \frac{dn^0_{k,\omega}}{d\tilde{q}_j}(1 - e^{-\lambda\bar{z}_{0,\omega}}) - n^0_{k,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = 0 \\
\frac{d^2\mathcal{L}}{d\varphi_{k,\omega}d\tilde{q}_j} &= \frac{dn^0_{k,\omega}}{d\tilde{q}_j}\theta_{k0}e^{-\lambda z^{L_k}_{k,\omega}} - n^0_{k,\omega}\theta_{k0}\lambda e^{-\lambda z^{L_k}_{k,\omega}}\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} - \frac{ds_{k,\omega}}{d\tilde{q}_j} = 0
\end{aligned}
$$

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

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

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

$f_{k,\omega}(\varphi_{0,\omega}) = 1$ if $L_k=0$, $f_{k,\omega}(\varphi_{0,\omega}) = 1-\theta_{kk}e^{-\lambda z^0_{k,\omega}}$ if $L_k=1$ and $f_{k,\omega}(\varphi_{0,\omega}) = 1-e^{-\lambda z^1_{k,\omega}}/\left[e^{-\lambda z^0_{k,\omega}}\left(1-\theta_{kk}e^{-\lambda z^0_{k,\omega}}\right)\right]$ if $L_k=2$. $f_{k,\omega}(\varphi_{0,\omega})>0$ (see section C.2.2).

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

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

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

$$\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} = \frac{\frac{\varphi_{0,\omega}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}} + \xi_{j,\omega}\sum_{k=0}^{1}\frac{s_{k,\omega}}{f_{k,\omega}(\varphi_{0,\omega})}}{\lambda(1+e^{-\lambda\bar{z}_{0,\omega}})\sum_{k=0}^{1}\frac{s_{k,\omega}}{f_{k,\omega}(\varphi_{0,\omega})}\sum_{k=0}^{1}\xi_{k,\omega}n^0_{k,\omega} + \varphi_{0,\omega}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}} > 0 \qquad \square$$

To show: Higher local output $\tilde{q}_j$ increases (decreases) the number of production workers at the same (other) location $n^0_{j,\omega}$ ($n^0_{k,\omega}, k \neq j$).
From $\frac{d^2\mathcal{L}}{d\xi_{k,\omega}d\tilde{q}_j}$ $j,k=0,1$:

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

By $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $\frac{dq_{k,\omega}}{d\tilde{q}_j}=0$, so $\frac{dn^0_{k,\omega}}{d\tilde{q}_j}<0$. $\frac{dq_{j,\omega}}{d\tilde{q}_j}=1$ and $\frac{d\bar{z}_{0,\omega}}{d\tilde{q}_j} < \frac{1}{n^0_{j,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}$, so $\frac{dn^0_{k,\omega}}{d\tilde{q}_j}<0.$ $\square$

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

From above:

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

This is the case if

$$\sum_{k=0}^{1}\xi_{k,\omega}n^0_{k,\omega}(1+e^{\lambda\bar{z}_{0,\omega}}) > \xi_{j,\omega}\frac{1}{\theta_{j0}}e^{\lambda z^{L_j}_{j,\omega}},$$

which holds whenever $s_{j,\omega} \geq 1/1+e^{\lambda\bar{z}_{0,\omega}}$ (sufficient, not necessary).

The positive impact of higher local output on knowledge follows from:

$$\frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} \overset{L_k=0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\tilde{q}_j} \qquad \frac{dz^{L_k}_{k,\omega}}{d\tilde{q}_j} - \frac{dz^{L_k-1}_{k,\omega}}{d\tilde{q}_j} \overset{L_k>0}{=} \frac{1}{\varphi_{0,\omega}\lambda}\frac{d\varphi_{0,\omega}}{d\tilde{q}_j}$$

C.3.4 Proposition 4: Comparative statics with respect to θ_{10}

We substitute $\frac{d\mathcal{L}}{dz^{\ell}_{j,\omega}}$, $\ell \leq L_j$, into $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\theta_{10}}$. We assume that $\xi_{j,\omega} \neq \tau\xi_{k,\omega}$, $j \neq k$ and $q_{j,\omega}$, $s_{j,\omega} > 0\,\forall j$. The second order conditions are, with $j \in \{0,1\}$:

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

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

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

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

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

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

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

$$\frac{d^2\mathcal{L}}{dz^{L_0}_{0,\omega}d\theta_{10}}\begin{cases}\overset{L_0=0}{=} -\frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{00}\lambda e^{-\lambda z^0_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^0_{0,\omega}}\frac{dz^0_{0,\omega}}{d\theta_{10}} = 0\\ \overset{L_0>0}{=} -w_0 c\lambda\theta_{00}e^{-\lambda z^{L_0-1}_{0,\omega}}\frac{dz^{L_0-1}_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{0,\omega}}{d\theta_{10}}\lambda\theta_{00}e^{-\lambda z^{L_0}_{0,\omega}} + \varphi_{0,\omega}\theta_{00}\lambda^2 e^{-\lambda z^{L_0}_{0,\omega}}\frac{dz^{L_0}_{0,\omega}}{d\theta_{10}} = 0\end{cases}$$

$$\frac{d^2\mathcal{L}}{dz^{L_1}_{1,\omega}d\theta_{10}}\begin{cases}\overset{L_1=0}{=} -\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}\lambda e^{-\lambda z^0_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^0_{1,\omega}}\frac{dz^0_{1,\omega}}{d\theta_{10}} - \varphi_{1,\omega}\lambda e^{-\lambda z^0_{1,\omega}} = 0\\ \overset{L_0>0}{=} -w_1 c\lambda\theta_{11}e^{-\lambda z^{L_1-1}_{1,\omega}}\frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}}\lambda\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}}\\ -\varphi_{1,\omega}\lambda e^{-\lambda z^{L_1}_{1,\omega}} = 0\end{cases}$$

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

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} if $L_1 \leq 1$ or if $L_1 = 2$ the establishment's share of CEO time is sufficiently high $s_{1,\omega} \geq 1 - f_1(\varphi_{0,\omega})$.

1. $\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{ds_{j,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\theta_{10}}$, $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$ 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 $L_j = 0$, $f_{j,\omega}(\varphi_{0,\omega}) = 1-\theta_{jj}e^{-\lambda z^0_{j,\omega}}$ if $L_j = 1$ and $f_{j,\omega}(\varphi_{0,\omega}) = 1 - e^{-\lambda z^1_{j,\omega}}/\left[e^{-\lambda z^0_{j,\omega}}\left(1-\theta_{jj}e^{-\lambda z^0_{j,\omega}}\right)\right]$ if $L_j = 2$.

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

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

3. Inserting $\frac{d\xi_{j,\omega}}{d\theta_{10}}$ and $\frac{dn^0_{j,\omega}}{d\theta_{10}}$ 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:

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

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 $\theta_{11}e^{-\lambda z^0_{1,\omega}} \leq 0.5 \leq 1-\theta_{11}e^{-\lambda z^0_{1,\omega}}$, 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 (b): The marginal benefit of CEO time $\varphi_{0,\omega}$ decreases with the helping costs θ_{10} if $L_1 \leq 1$ or the establishment's share of CEO time is sufficiently high $s_{1,\omega} \geq 1-f_1(\varphi_{0,\omega})$.
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 (a, b): The knowledge of the employees at all below-CEO layers $z^\ell_{j,\omega}$, $\ell \leq L_j$ increases with the helping costs θ_{10} at the establishment and 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 $s_{1,\omega} \geq 1-f_1(\varphi_{0,\omega})$.
$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\theta_{10}}$ yields:

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

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

$\frac{dz^\ell_{0,\omega}}{d\theta_{10}} < 0$ 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 (a, b): 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 $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$. In result, the total number of production workers decreases. □

To show (a, b): The marginal production costs $\xi_{j,\omega}$ increases with the helping costs θ_{10} at the establishment and decreases with the helping costs θ_{10} at the headquarters.

- $j = 1$: Follows because $\varphi_{1,\omega}e^{-\lambda z_{1,\omega}^{L_1}} > \xi_{j,\omega}\lambda e^{-\lambda \bar{z}_{0,\omega}}\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{0,\omega}}{d\theta_{10}}\theta_{j0}e^{-\lambda z_{j,\omega}^{Lj}}$, which results after substituting for $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}$ and $\frac{d\varphi_{0,\omega}}{d\theta_{10}}$ because $s_{1,\omega}\frac{1-f_{1,\omega}(\varphi_{0,\omega})}{f_{1,\omega}(\varphi_{0,\omega})} < \sum_{j=0}^{1}\frac{s_{j,\omega}}{f_{j,\omega}(\varphi_{0,\omega})}$.
- $j = 0$: 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 $s_{1,\omega} > s_{1,\omega}(1 - f_{1,\omega}(\varphi_{0,\omega}))$. □

C.3.5 Proposition 5: The optimal number of layers

Parameter conditions: $w_1 = w_0$, $\theta_{10} = \theta_{00}$, $\tau = 1$.

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

The firm with the L_0/L_0-organization chooses the same knowledge levels in the headquarters and the establishment 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 is equal to the cost function of a single establishment firm with $n^0_{0,L_0+1} = \sum_{j=0}^{1} n^0_{j,L_0/L_0}$, so Proposition 1b) applies.

b) To show): The average cost of the $L_0/L_0 + 1$-organization and the $L_0 + 1/L_0$-organization coincide.

Follows from $w_1 = w_0$, $\theta_{10} = \theta_{00}$.

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

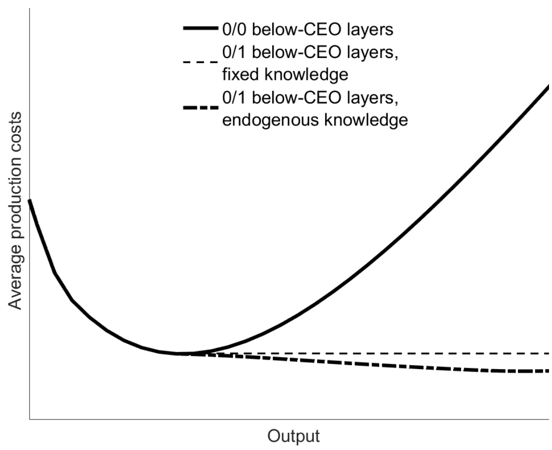
To show: The average cost function of the $L_0/L_0 + 1$-organization decreases with output $\tilde{q}$ for $\tilde{q}^*_{L_0+1/L_0+1} > \tilde{q} > \tilde{q}^*_{L_0/L_0}$.

For simplicity and without loss of generality, we choose $L_0 = 0$. The proof proceeds in two steps. First, we construct a 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 $\tilde{q} \in [\tilde{q}^*_{0/0}, \tilde{q}^{MAX}]$ with constant costs. Second, we show that the average cost function of the organization with 0/1 below-CEO layers and endogenous knowledge levels decreases with output for $\tilde{q} \in [\tilde{q}^*_{0/0}, \tilde{q}^*_{1/1})$, so it is lower than the average cost function of the organization with fixed knowledge levels. Figure C.1 illustrates the argument.

1. We construct an ME organization with $L_0 = 0$ below-CEO layers at the headquarters and $L_1 = 1$ below-CEO layer at the establishment that has the same average cost as an ME organization with $L_1 = L_0 = 0$ below-CEO layers at the minimum efficient scale $\tilde{q}^*_{0/0}$.

   The knowledge levels of the 0/0-organization coincide with the knowledge levels of a single establishment firm with no below-CEO layer (i.e., $L = 0$). Thus, at

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



The figure illustrates part b) of the proof of Proposition 5. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{10} = \theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = w_1 = 1$. The solid line refers to an organization with 0/0 below-CEO layers. The dashed lines show the average cost functions of organizations with 0/1 below-CEO layers. The light line refers to the organization with fixed knowledge levels, the bold line to the organization with endogenous knowledge levels.

the minimum efficient scale $\tilde{q}^*_{0/0}$,

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

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

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

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

Fix the knowledge levels of an ME firm with organization $\omega = 0/1$ such that

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

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

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

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

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

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

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

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

$$\xi_{1,0/1} = \xi_{0,0/1} = \xi_{0,0} \quad \text{by construction} \tag{C.10}$$

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

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

The ME firm produces output $\tilde{q}$ with $\tilde{q}^{MAX} \geq \tilde{q} \geq \tilde{q}^*_{0/0}$ by allocating the share s of output to the headquarters and the share $1-s$ to the establishment, where

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

The numerator and denominator are negative. The denominator is constant. $0 \leq s \leq 1$, because the numerator achieves its minimum at $\tilde{q} = \frac{1}{\theta_{00}} e^{\lambda z^{0\,MES}_{0,0}} (1 - e^{-\lambda \bar{z}^{MES}_{0,0}})$ $(s = 1)$, and its maximum at $\tilde{q} = \frac{1}{\theta_{00}} e^{\lambda z^1_{1,0/1}} (1 - e^{-\lambda \bar{z}^{MES}_{0,0}})$ $(s = 0)$. That is, the average cost function of the ME firm with fixed knowledge levels is flat for $\tilde{q} \in \left[\tilde{q}^*_{0/0}, \tilde{q}^{MAX}\right]$ (see the light dashed line in Figure C.1). □

2. We show that the average cost function of 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 firm with 0/0 below-CEO layers 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

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

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

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

   $\varphi_{j,0/1}$ is constant; $\bar{z}_{0,0/1}$ increases with $\tilde{q}$ by Proposition C.1. The maximum value of $AC_{0,0/1}(\tilde{q})$ is $AC(\tilde{q}, \bar{z}^{MES}_{0,1}, z^{0\,MES}_{0,1}, z^0_{1,0/1}(z^{0\,MES}_{0,1}), z^1_{1,0/1}(z^{0\,MES}_{0,1})) = AC^{MES}_{0,0}$. At $\tilde{q}^*_{0/0}$, $\xi_{0,0/1} = AC_{0,0}$; $\xi_{0,0/1}$ decreases with $\tilde{q}$ for $\tilde{q}^{1/1}_{0/1} > \tilde{q} > \tilde{q}^*_{0/0}$. The average cost function $AC_{0,1/1}(\tilde{q})$ decreases for $\tilde{q}^*_{1/1} > \tilde{q} \geq \tilde{q}^{1/1}_{0/1}$ 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 $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$ 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 output $\tilde{q}^{L_0+1/L_0+1}_{L_0/L_0+1} > \tilde{q}^{L_0+1/L_0+1}_{L_0/L_0}$.

We exploit the characteristics of the average cost function.

- $AC_{0,0/1} \leq AC^{MES}_{0,0} \forall \tilde{q}^*_{0/0} \leq \tilde{q} \leq \tilde{q}^{MAX}$;
- $AC_{0,0/0}$ is increasing for $\tilde{q} > \tilde{q}^*_{0/0}$;
- $AC_{0,1/1}$ is decreasing for $\tilde{q} \leq \tilde{q}^*_{1/1}$, where $\tilde{q}^{MAX} \leq \tilde{q}^*_{1/1}$;
- at $\tilde{q}^*_{0/0}$, $AC_{0,1/1} > AC_{0,0/0}$.

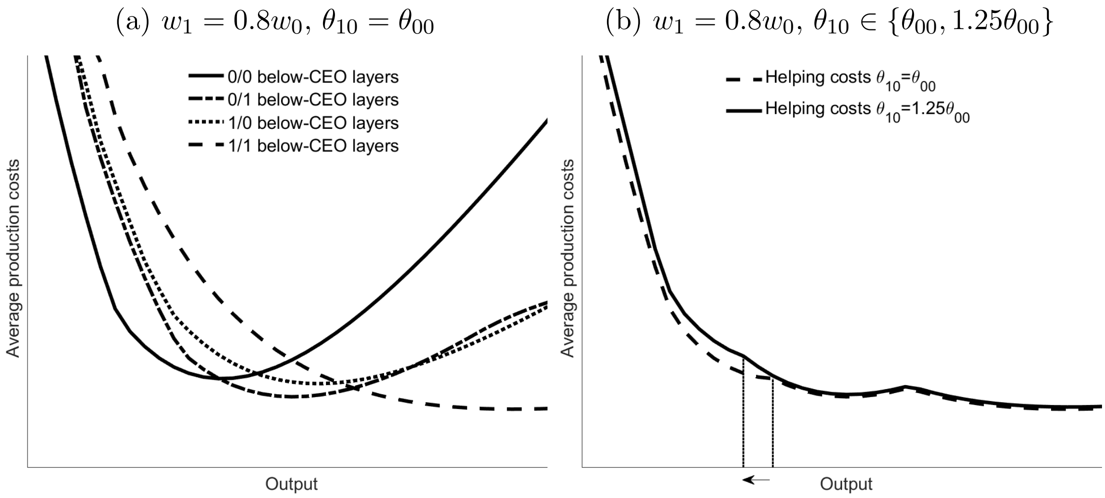
In consequence, the increasing average costs function of the ME firm with 0/0 below-CEO layers $AC_{0,0/0}$ intersects the decreasing average costs function of the ME firm with 1/1 below CEO layers $AC_{0,1/1}$ at a lower output than the level at which the decreasing average cost function of the ME firm with 0/1 below-CEO layers $AC_{0,0/1}$ intersects the average cost function $AC_{0,1/1}$. □

C.3.6 The optimal number of layers with transport frictions, wage and output differences

Wage differences, $w_1 \neq w_0$, $\tau > 1$, $\theta_{10} \geq \theta_{00}$ and $\tilde{q}_1 = \tilde{q}_0$.

Note: Output is on a log scale. The decrease of output in Figure C.3c compared to the decrease in Figure C.2b is larger than the lengths of the arrows suggest.

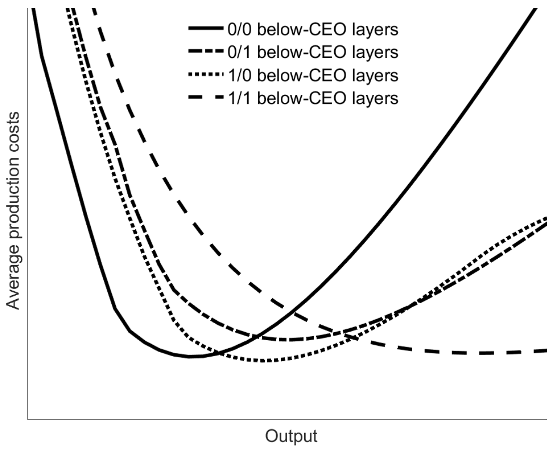
Figure C.2: Lower wages at the establishment than at the headquarters



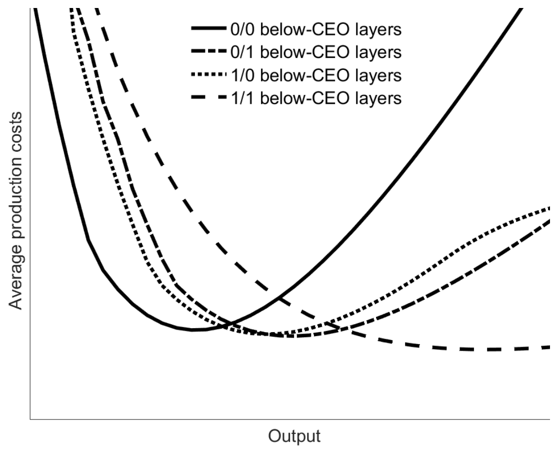
The figure plots the average cost functions of a ME firm for $w_1 = 0.8w_0$, $\tau = 1.5$, and $\tilde{q}_1 = \tilde{q}_0$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$, $\tilde{q}_j \in [1, 63]$.
(a): $\theta_{10} = \theta_{00}$: Due to the lower wages at the establishment, the firm adds a layer of middle managers first only at the establishment and then also at the headquarters.
(b): $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$: Higher helping costs across space decrease the level of output at which the firm adds a layer of middle managers at the establishment.
Note: for the highest values of output in Figure (a), a layer of middle managers only at the headquarters has lower average costs than a layer only at the establishment. The reason is that a firm with 0/1 organization shifts total production to the establishment for high values of output and thus bears the transport costs, while it sticks to multi-establishment production with the 1/0 organization. The costs of the 0/1 organization are below those of the 1/0 organization for even higher output.

Figure C.3: Higher wages at the establishment than at the headquarters

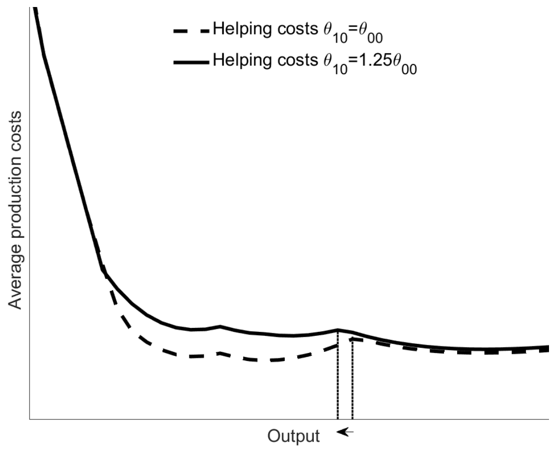
(a) $w_1 = 1.25w_0$, $\theta_{10} = \theta_{00}$



(b) $w_1 = 1.25w_0$, $\theta_{10} = 1.25\theta_{00}$



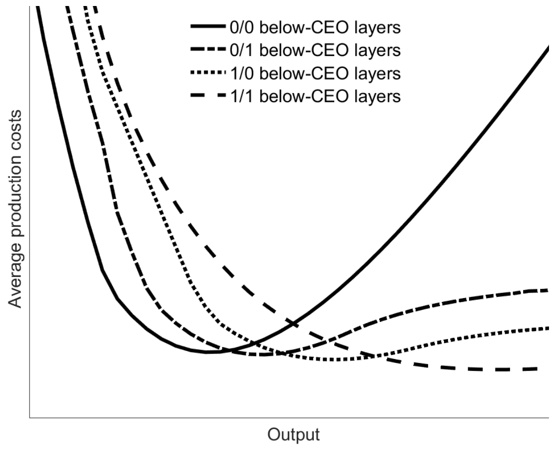
(c) $w_1 = 1.25w_0$, $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$



The figure plots the average cost functions of a ME firm for $w_1 = 1.25w_0$, $\tau = 1.5$, and $\tilde{q}_1 = \tilde{q}_0$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$, $\tilde{q}_j \in [1, 63]$.
(a): $\theta_{10} = \theta_{00}$: Due to the lower wages at the headquarters, the firm adds a layer of middle managers first only at the headquarters and then also at the establishment.
(b): $\theta_{10} = 1.25\theta_{00}$: The helping costs across space outweigh the higher wages at the establishment, so the firm adds a layer of middle managers only at the establishment for a range of output levels.
(c): $\theta_{10} \in \{\theta_{00}, 1.25\theta_{00}\}$: Higher helping costs across space decrease the level of output at which the firm adds a layer of middle managers at the headquarters in addition to the establishment.
Note: for the highest values of output in the figure, adding a layer only at the establishment has lower average costs than adding it only at the headquarters. As in Figure C.2a, the reason is that the firm shifts total production to the headquarters with the 1/0 organization for high values of output and thus bears the transport costs.

Output differences, $\tilde{q}_1 \neq \tilde{q}_0$, $\tau > 1$, $\theta_{10} = \theta_{00}$ and $w_1 = w_0$.

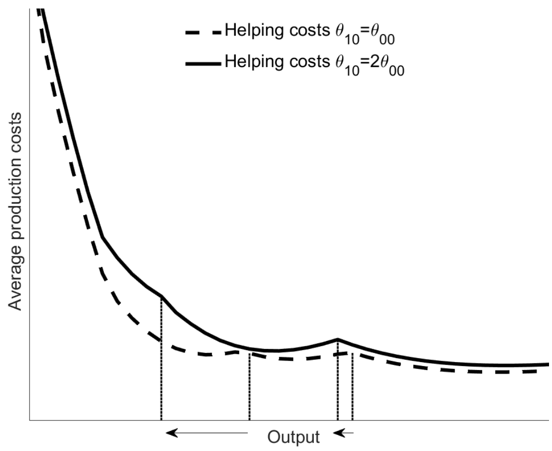
Figure C.4: Lower output at the establishment than at the headquarters



The figure plots the average cost functions of a ME firm for $\tilde{q}_1 = 0.5\tilde{q}_0$, $w_1 = w_0$, $\tau = 1.1$, and $\theta_{10} = \theta_{00}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$, $\sum_{j=0}^{1} \tilde{q}_j \in [2, 126]$.

Helping cost and output differences, $\tau > 1$ and $w_1 = w_0$.
Note: Output is on a log scale. The relative decrease of output in levels therefore differs from what the relative length of the arrows suggests. In levels, the decrease of output for an additional layer at the headquarters is half the decrease at the establishment.

Figure C.5: Simultaneous change of helping costs and output



The figure plots the minimum average cost functions of a ME firm for $w_1 = w_0$, $\tau = 1.1$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$, $\sum_{j=0}^{1} \tilde{q}_j \in [2, 126]$. The dashed line assumes $\theta_{10} = \theta_{00}$ and $\tilde{q}_1 = \tilde{q}_0$. The solid line assumes $\theta_{10} = 2\theta_{00}$ and $\tilde{q}_1 = .66\tilde{q}_0$.

C.3.7 Comparative statics with respect to $\tilde{q}_j$, θ_{10} if $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $\tau \geq 1$

Proposition C.1. *Suppose the firm produces in the headquarters and the establishment. Suppose either that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other and that $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, or that there are no transport frictions $\tau = 1$, so $\xi_{0,\omega} = \xi_{1,\omega}$, but the headquarters and the establishment are not symmetric, i.e., $\theta_{10} \geq \theta_{00}$, and $w_1 < w_0$ or $L_1 \neq L_0$. Given the organizational structure ω,* *a) CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_j$. Higher local output $\tilde{q}_j$ increases the number of production workers $n^0_{k,\omega}$ at the location with the larger decrease of the marginal production costs and decreases their number at the other location, unless $L_0 > L_1$ and wages w_1 are too high.*

b) The knowledge of the employees at all below-CEO layers $z^\ell_{k,\omega}$, $\ell \leq L_k$ and the marginal benefit of CEO time $\varphi_{k,\omega}$, $k = 0, 1$, do not vary with local output $\tilde{q}_j$.

c) The marginal production costs $\xi_{k,\omega}$, $k = 0, 1$, decrease with output $\tilde{q}_j$.

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

Intuition.

- If $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, the multi-establishment firm effectively produces with two different technologies that have the same effective marginal production costs. It grows by recombining the technologies through reallocating CEO time and output.
- CEO knowledge increases with local output, because knowledge and labor are complementary inputs. The number of production workers increases (decreases) at the location with the larger (smaller) decrease of the marginal costs, because the firm reallocates output to the location with the larger decrease of the marginal costs.
- The parameter values determine the optimal combination of the knowledge levels of the employees at all below-CEO layers. The optimal combination is uniquely given by the two conditions $\varphi_{0,\omega} = \varphi_{1,\omega}$ and $\xi_{0,\omega} = \xi_{1,\omega}$. Correspondingly, neither knowledge levels nor the marginal benefit of CEO time vary with local output.

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$, $\tau \geq 1$. The second order conditions correspond to the ones in Appendix section C.3.3 with the following exceptions:

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

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_0} = 1 - \frac{dq_{0,\omega}}{d\tilde{q}_0} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_0} = 0$$

$$\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\tilde{q}_1} = \tau - \frac{dq_{0,\omega}}{d\tilde{q}_1} - \tau\frac{dq_{1,\omega}}{d\tilde{q}_1} = 0$$

To show (a): CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_0$, $\tilde{q}_1$ (or, if $\xi_{0,\omega} = \xi_{1,\omega}$, total output $\tilde{q}$).

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

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

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

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

Case 2: $\xi_{0,\omega} = \tau\xi_{1,\omega}$. The second order conditions correspond to the ones in Appendix section C.3.3 with the following exceptions:

$$\frac{d^2\mathcal{L}}{dq_{0,\omega}d\tilde{q}_j} - \frac{d^2\mathcal{L}}{dq_{1,\omega}d\tilde{q}_j} = \frac{d\xi_{0,\omega}}{d\tilde{q}_j} - \tau\frac{d\xi_{1,\omega}}{d\tilde{q}_j} = 0$$
$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_0} = \tau - \tau\frac{dq_{0,\omega}}{d\tilde{q}_0} - \frac{dq_{1,\omega}}{d\tilde{q}_0} = 0$$
$$\frac{d^2\mathcal{L}}{d\underline{\phi}_{0,\omega}d\tilde{q}_1} = 1 - \tau\frac{dq_{0,\omega}}{d\tilde{q}_1} - \frac{dq_{1,\omega}}{d\tilde{q}_1} = 0$$

To show (a): CEO knowledge $\bar{z}_{0,\omega}$ increases with local output $\tilde{q}_0$, $\tilde{q}_1$.
By an analogous argument to Case 1:

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

To show (a): Higher local output $\tilde{q}_k$ increases the number of production workers $n^0_{j,\omega}$ at the location with the larger decrease of the marginal production costs (i.e., the headquarters) and decreases their number at the other location (i.e., the establishment), unless $L_0 > L_1$ and wages w_1 are too high.
Follows from an analogous argument to Case 1.

To show (b): 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 output $\tilde{q}_0$, $\tilde{q}_1$.
Follows from an analogous argument to Case 1.

To show (c): The marginal production cost $\xi_{k,\omega}$ decreases with local output $\tilde{q}_j$.
Follows from an analogous argument to Case 1.

Proposition C.2. *Suppose the firm produces in the headquarters and the establishment and that $\theta_{10} > \theta_{00}$. Suppose either that the firm incurs transport costs $\tau > 1$ to ship output from one location to the other and that $\xi_{j,\omega} = \tau\xi_{k,\omega}$, $j \neq k$, or that there are no transport frictions $\tau = 1$, so $\xi_{0,\omega} = \xi_{1,\omega}$, but the headquarters and the establishment are not symmetric, i.e., $w_1 < w_0$ or $L_1 \neq L_0$. Given the organizational structure ω,*

a) CEO knowledge $\bar{z}_{0,\omega}$, the marginal benefit of CEO time $\varphi_{j,\omega}$ and the knowledge of the employees at all below-CEO layers $z^\ell_{j,\omega}$, $\forall \ell < L_j$, increase with the helping costs θ_{10}.

b) 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. The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n^\ell_{j,\omega}$, $\forall \ell < L_j$, decreases.

c) The marginal production cost $\xi_{j,\omega}$ increase with the helping costs θ_{10}.

The comparative statics hold if $L_0 \leq L_1$, or $L_0 > L_1$ and wages w_1 are sufficiently small.

Intuition.

- If the marginal costs including transport costs are equal at the headquarters and the establishment, the firm reallocates output from the establishment to the headquarters in response to higher θ_{10}. In consequence, the number of production workers at the headquarters increases, as do their knowledge, the marginal benefit of CEO time and the marginal production costs.

- The intuition for the increase of CEO knowledge, the increase of the knowledge and the decrease of the number of employees at the establishment is analogous to Proposition 4.

Case 1: $\xi_{1,\omega} = \tau\xi_{0,\omega}$. The second order conditions correspond to the ones in Appendix section C.3.4 with the following exceptions:

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

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 $\tau\frac{d\xi_{0,\omega}}{d\theta_{10}} - \frac{d\xi_{1,\omega}}{d\theta_{10}} = 0$, $\tau\xi_{0,\omega} = \xi_{1,\omega}$, $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$ 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 $L_0 \leq L_1$ and for $L_0 > L_1$ unless wages w_1 are too high, see above $({dn^\ell_{k,\omega}}/{d\tilde{q}_j})$.

To show (a): The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with the helping costs θ_{10}.
Follows from $\frac{d\varphi_{1,\omega}}{d\theta_{10}} = \frac{d\varphi_{0,\omega}}{d\theta_{10}} > 0$ 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}.

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 (b): The number of production workers at the establishment $n^0_{1,\omega}$ decreases and the number of production workers at the headquarters $n^0_{0,\omega}$ increases with the helping costs θ_{10}.

Substituting $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$ into $\frac{d^2\mathcal{L}}{d\bar{\phi}_{0,\omega}d\theta_{10}}$ yields:

$$\frac{dn^0_{1,\omega}}{d\theta_{10}} = -\frac{1}{\tau}\frac{dn^0_{0,\omega}}{d\theta_{10}} - \frac{1}{\tau}\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}(n^0_{0,\omega} + \tau n^0_{1,\omega})\frac{d\bar{z}_{0,\omega}}{d\theta_{10}}$$

Substituting this expression into $\frac{d^2\mathcal{L}}{d\bar{\kappa}_{0,\omega}d\theta_{10}}$ with $\frac{d^2\mathcal{L}}{d\varphi_{j,\omega}d\theta_{10}}$ yields:

$$\frac{dn^0_{0,\omega}}{d\theta_{10}} = \frac{e^{\lambda z^{L_0}_{0,\omega}}e^{\lambda z^{L_1}_{1,\omega}}}{\tau\theta_{00}e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10}e^{\lambda z^{L_0}_{0,\omega}}} \times \left(\frac{\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1-e^{-\lambda\bar{z}_{0,\omega}}}(n^0_{0,\omega} + \tau n^0_{1,\omega})\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} + \tau\sum_{j=0}^{1}\lambda s_{j,\omega}\frac{dz^{L_j}_{j,\omega}}{d\theta_{10}} - \tau n^0_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}} \right) > 0$$

by $\lambda s_{1,\omega}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} > n^0_{1,\omega}e^{-\lambda z^{L_1}_{1,\omega}}$ as $\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}} > \frac{1}{\lambda\theta_{10}}$.

To show (b): 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 (c): 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.3.4 with the following exceptions:

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

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

Results follow from derivations analogous to those for Case 1.

C.4 Proposition 6: The optimal output

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

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

$$\frac{\partial \pi_i}{\partial \tilde{q}_j} = \frac{\partial p_j}{\partial \tilde{q}_j}\tilde{q}_j + p_j(\tilde{q}_j) - \xi_{j,\omega} = 0$$

$\tau\xi_{j,\omega} \neq \xi_{k,\omega}$. We define $\hat{q}_j \equiv -\tilde{q}_j$. From Proposition 4:

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

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

By monotone comparative statics, $\hat{q}_0$ and $\tilde{q}_1$ decrease with the helping costs θ_{10} if

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

This holds for sufficiently high output $\tilde{q}_j$. In result, $\tilde{q}_0$ increases and $\tilde{q}_1$ decreases with the helping costs θ_{10}. □

$\tau\xi_{j,\omega} = \xi_{k,\omega}$. From Proposition C.2:

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

From Proposition C.1:

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

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

D Reorganization due to high-speed train routes

D.1 Travel time data

Table D.1: Reduction of travel times in minutes

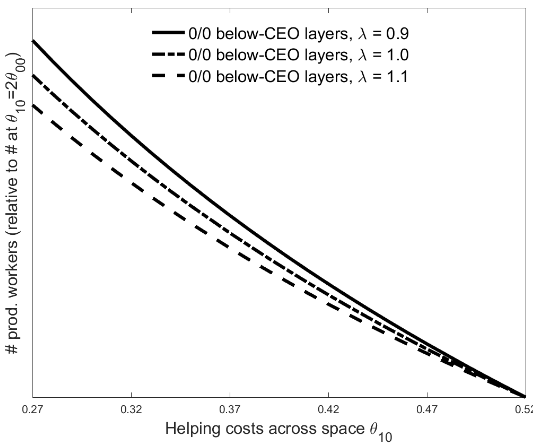
	High speed	Mean	p25	p50	p75
2000-2004	0	-1.6	-5.8	0.2	5.1
	1	-22.7	-51.5	-8.7	3.6
2004-2008	0	-1.4	-5.8	-0.2	3.1
	1	-16.8	-28.8	-9.9	-1.2

The table displays summary statistics on the reduction of travel time between 2000 and 2004 and 2004 and 2008 separately for connections where at least 50% of passengers use the new high speed routes and other routes.

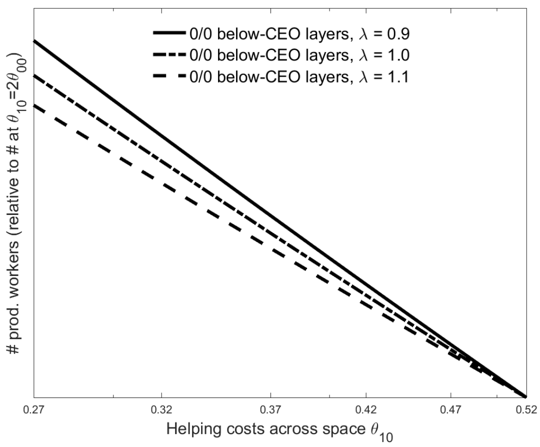
D.2 Effect heterogeneity depending on the predictability of the production process λ

Figure D.1: Heterogeneity of the effect of θ_{10} depending on the predictability λ

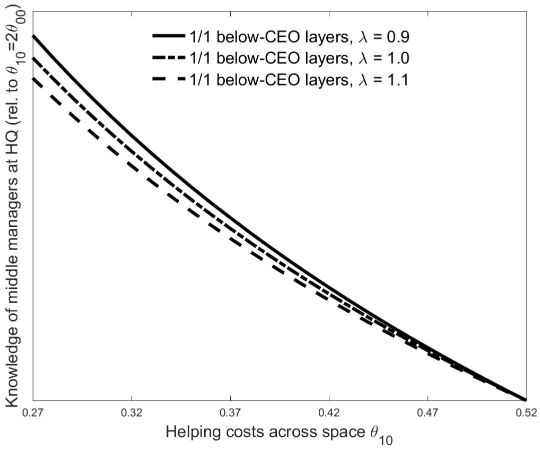
(a) # workers, headquarters, $\omega = 0/0$



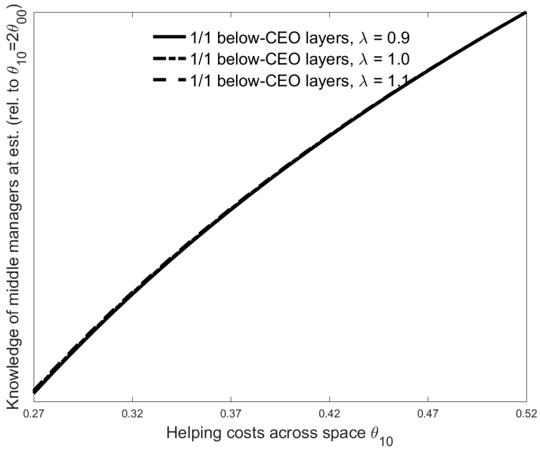
(b) # workers, establishment, $\omega = 0/0$



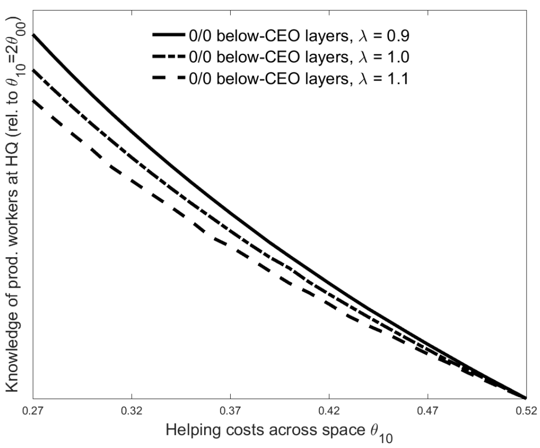
(c) Mg. knowledge, headquarters, $\omega = 1/1$



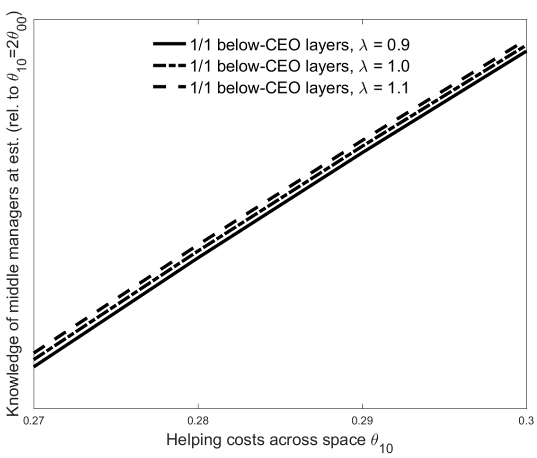
(d) Mg. knowledge, establishment, $\omega = 1/1$



(e) Mg. knowledge, headquarters, $\omega = 0/0$



(f) Close-up: Mg. knowledge, est., $\omega = 1/1$



The figure plots the number and knowledge of production workers and middle managers of a ME firm as a function of the helping costs across space $\theta_{10} \in [.27, .52]$ for different values of λ with $\omega \in \{0/0, 1/1\}$, $\tau = 1.1$, $w_1 = w_0$, and $\tilde{q}_1 = \tilde{q}_0 = 50$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{00} = .26$ (Caliendo and Rossi-Hansberg, 2012), $w_0 = 1$. As the effect of λ relative to the effect of θ_{10} is small for the managerial knowledge at the establishment, Figure D.1f provides a close-up of the simulation for small values of θ_{10}.

D.3 Robustness checks

Table D.2: Regression results, mean travel times, 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	0.093***	0.016*	0.245	0.114	0.097***	0.015+	0.498*	0.192
	(0.018)	(0.008)	(0.218)	(0.139)	(0.019)	(0.008)	(0.238)	(0.151)
# 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
R-squared	0.802	0.874	0.912	0.868	0.803	0.878	0.911	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.020	−0.008	0.602*	0.007	0.004	0.000	0.995**	0.562*
	(0.017)	(0.014)	(0.247)	(0.191)	(0.021)	(0.018)	(0.311)	(0.243)
# 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
R-squared	0.909	0.883	0.935	0.892	0.932	0.889	0.935	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					0.037+	0.024**	0.908***	0.456**
					(0.019)	(0.008)	(0.233)	(0.149)
# observations					70,705	70,705	70,705	70,705
# est.					10,835	10,835	10,835	10,835
R-squared					0.808	0.882	0.913	0.871
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Robust standard errors in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. All variables are winsorized at the first and 99th percentiles. Treatment dummies are equal to one if the mean travel time between the headquarters and the establishment decreases by at least 30 minutes.

Table D.3: Regression results, establishment-level layer definition, 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)
# 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
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
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)
# 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
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
Non-directly affected establishment								
Firm treated					−0.030**	0.004	0.221	0.412**
					(0.011)	(0.008)	(0.235)	(0.140)
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995
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

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. Managerial outcomes in columns 2, 3, 6 and 7 are defined treating the lowest-level layer in each establishment/the HQ as non-managerial. All variables are winsorized at the first and 99th percentiles.

Table D.4: Regression results, Bundesland-year fixed effects, 2000-2010 panel

	All firms				Firms with ≥ 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	0.077^{***}	0.014^{+}	-0.145	-0.210	0.082^{***}	0.011	-0.084	-0.229
	(0.016)	(0.008)	(0.205)	(0.132)	(0.017)	(0.008)	(0.218)	(0.140)
# 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
R-squared	0.800	0.873	0.911	0.866	0.800	0.877	0.910	0.867
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
Bula-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	-0.016	0.029^{*}	0.820^{***}	0.211	-0.031^{+}	0.053^{**}	1.347^{***}	0.804^{***}
	(0.016)	(0.013)	(0.238)	(0.178)	(0.018)	(0.017)	(0.292)	(0.222)
# 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
R-squared	0.905	0.877	0.931	0.887	0.924	0.879	0.930	0.890
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ b.-year FE	Y	Y	Y	Y	Y	Y	Y	Y

Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. *Bula-year FE* (*HQ b.-year FE)*: fixed effects at (headquarter) Bundesland $\times$ year level. All variables are winsorized at the first and 99th percentiles.

Table D.5: Regression results, cluster robust SE, 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	0.061**	0.004	−0.213	−0.145	0.074*	0.001	−0.033	−0.147
	(0.023)	(0.012)	(0.325)	(0.213)	(0.029)	(0.014)	(0.334)	(0.228)
# 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
R-squared	0.802	0.874	0.912	0.868	0.803	0.878	0.911	0.869
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.015	0.019	0.688	0.063	−0.018	0.041+	1.054	0.631
	(0.020)	(0.019)	(0.491)	(0.336)	(0.026)	(0.022)	(0.647)	(0.504)
# 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
R-squared	0.909	0.883	0.935	0.892	0.932	0.890	0.935	0.897
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.003	0.009	0.523+	0.412+
					(0.029)	(0.012)	(0.296)	(0.223)
# observations					72,040	72,040	72,040	72,040
# est.					10,995	10,995	10,995	10,995
R-squared					0.807	0.883	0.912	0.873
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Standard errors clustered by establishment (headquarters) and establishment (headquarters) county in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. All variables are winsorized at the first and 99th percentiles. P-values for headquarters: 16.3% in column 3, 10.6% in column 7.

Table D.6: Regression results, without moving establishments/ headquarters, 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	0.064***	0.007	−0.057	−0.106	0.082***	0.002	0.031	−0.177
	(0.018)	(0.009)	(0.232)	(0.149)	(0.019)	(0.009)	(0.246)	(0.160)
# observations	88,873	88,873	88,873	88,873	79,363	79,363	79,363	79,363
# est.	12,861	12,861	12,861	12,861	11,658	11,658	11,658	11,658
R-squared	0.804	0.878	0.914	0.871	0.804	0.881	0.912	0.872
Est. FE	Y	Y	Y	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Headquarters								
Firm treated	−0.014	0.014	0.719**	−0.012	−0.019	0.050**	1.179***	0.587*
	(0.018)	(0.015)	(0.250)	(0.198)	(0.021)	(0.019)	(0.308)	(0.259)
# observations	21,719	21,719	21,719	21,719	11,633	11,633	11,633	11,633
# HQ	2,726	2,726	2,726	2,726	1,502	1,502	1,502	1,502
R-squared	0.916	0.889	0.939	0.895	0.935	0.893	0.939	0.900
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					−0.013	0.008	0.441*	0.221
					(0.019)	(0.008)	(0.223)	(0.143)
# observations					68,192	68,192	68,192	68,192
# est.					10,483	10,483	10,483	10,483
R-squared					0.808	0.886	0.913	0.877
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. All variables are winsorized at the first and 99th percentiles. Sample excludes establishments (headquarters) that move between counties.

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

	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	0.079***	−0.001	−0.152	−0.311$^+$	0.095***	0.002	0.068	−0.296
	(0.022)	(0.011)	(0.283)	(0.185)	(0.023)	(0.011)	(0.301)	(0.197)
# observations	63,359	63,359	63,359	63,359	56,886	56,886	56,886	56,886
# est.	9,097	9,097	9,097	9,097	8,267	8,267	8,267	8,267
R-squared	0.798	0.871	0.911	0.870	0.799	0.875	0.911	0.872
Est. FE	Y	Y	Y	Y				
County-year FE	Y	Y	Y	Y				
Headquarters								
Firm treated	0.003	−0.000	0.970***	0.060	−0.030	0.011	1.400***	0.577*
	(0.021)	(0.017)	(0.288)	(0.220)	(0.025)	(0.021)	(0.369)	(0.283)
# observations	16,521	16,521	16,521	16,521	8,985	8,985	8,985	8,985
# HQ	2,063	2,063	2,063	2,063	1,559	1,559	1,559	1,559
R-squared	0.906	0.884	0.939	0.895	0.931	0.888	0.937	0.903
HQ FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ c.-year FE	Y	Y	Y	Y	Y	Y	Y	Y
Non-directly affected establishment								
Firm treated					0.000	0.005	0.726**	0.392*
					(0.022)	(0.009)	(0.257)	(0.168)
# observations					49,052	49,052	49,052	49,052
# est.					7,439	7,439	7,439	7,439
R-squared					0.805	0.882	0.914	0.877
Est. FE					Y	Y	Y	Y
County-year FE					Y	Y	Y	Y
HQ c.-year FE					Y	Y	Y	Y

Robust standard errors in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variables:* see Table 7. All variables are winsorized at the first and 99th percentiles. Sample is restricted to establishments (headquarters) that are connected to the long distance network in all years.