Geographic Frictions, ICT Access, and Firm Organization

Microeconometric Evidence from Administrative Data

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2018

vorgelegt von
Henrike Steimer

Referentin:	Prof. Dr. Monika Schnitzer
Koreferent:	Prof. Dr. Florian Englmaier
Promotionsabschlussberatung:	11. Juli 2018

Datum der mündlichen Prüfung: 25. Juni 2018

Namen der Berichterstatter: Monika Schnitzer, Florian Englmaier, Oliver Falck

Für meine Familie.

Acknowledgements

First and foremost, I would like to thank my supervisor Monika Schnitzer for her continuous and outstanding support over the past years. She has given me the opportunity to gain insights not only within economics, but also on a professional and on a personal level. I am indebted to her for much of what I take away from these years. Florian Englmaier also played a very important role. I thank him for always wholeheartedly supporting me and for his positive attitude. I would also like to thank Oliver Falck for completing my dissertation committee. He has always been very helpful and supportive.

I am very grateful to Anna Gumpert, without whom this dissertation would never have been possible. Her guidance, her patience with my ideas, and her example inspired me to push harder, learn more, and think deeper about economics.

I thank my colleagues at the Seminar for Comparative Economics. Their thoughtful advice and sincere support were instrumental in many moments during my studies. In addition, I thank them for the great times we had together. They made this time very special. My colleagues at the Munich Graduate School of Economics were always helpful and made the years of graduate school much nicer. Cathrin Mohr, Markus Nagler, Lisa Simon, and Matthias Wilhelm deserve special mention. They helped to make some difficult times less painful and the good times even better. I would also like to thank Dagmar Ehrhardt, Karin Fritsch, Florian Urbschat, Julia Zimmermann, and the student assistants at the Seminar for Comparative Economics for their great help.

My co-authors Manfred Antoni, Anna Gumpert, Eduardo Morales, and Ezra Oberfield deserve special credit for this dissertation. I learned a lot from working with them and would have not been able to realize any of the projects without the insights I gained from our joint work and the discussions. My "second office" during my studies was located at the Institute for Employment Research in Nuremberg. I thank the team at the research data centre for their hospitality. Florian Zimmermann provided excellent research assistance.

I would also like to thank the German Science Foundation for their financial support through Research Training Group 1928: Microeconomic Determinants of Labor Productivity and through CRC TR 190. This dissertation also benefitted from numerous comments and discussions at conferences and seminars.

Without my good friends, I would not have been able to achieve this. And more importantly, life would have been much less enjoyable. I am very grateful for having you. Most of all, I thank my family. Without their endless love and support, none of this would have been possible.

Henrike Steimer
July 2018

Contents

List of Tables

List of Figures

Preface

In today's world, information is an important resource. Information and communication technologies (ICT) have enabled a transformation of our society to the information age. They interrupted and accelerated the normal pace of economic progress in an economically significant manner. They are thus considered general purpose technologies (e.g. Brynjolfsson and McAffee, 2014). Prices to collect, store, and transmit information have fallen significantly. The number of transistors on integrated circuits, determining the processing power of electronic products, increased exponentially from a few thousand in the 1960s to several billions in the 2010s (e.g. Brynjolfsson and McAffee, 2014). The resulting increase in processing speed enabled the progress of the digital revolution.

Baldwin (2016) refers to ICT as the main driver of the second era of globalization by radically reducing the cost of moving ideas across borders. ICT facilitate the interaction and coordination between firms, customers, and suppliers (Brynjolfsson and Saunders, 2010). Since the introduction of broadband internet, e-mail and voice over internet protocol technologies offer cheap and convenient ways to communicate, replacing expensive airmail or overseas calls. ICT facilitate offshoring of business support services such as customer services, credit bureaus, and call centres. Cheap communication and easy access to remote servers further allow employees to work from distant locations to their employers (Bloom et al., 2015). Besides, firms set up e-businesses to reach new customers instead of sending catalogues. These new opportunities promote economic growth and efficiency.

Despite the many advantages of ICT, there is an increasing discussion about the unintended negative consequences of these new technologies. Many worry about increasing unemployment as humans are replaced by machines (e.g. Brynjolfsson and McAffee, 2011; Brynjolfsson and McAffee, 2014). Most tasks that are substituted by machines are routine in nature and executed by low-skilled individuals. Convincing evidence by many scholars shows that the ongoing technological change is skill-biased (e.g. Autor et al., 2003; Acemoglu and Autor, 2011, Acemoglu and Restrepo, 2017; on broadband internet: Akerman et al., 2015). Until now, most evidence on the impact of ICT on employment and skill composition is shown at the local labor market level. To draft effective policy, however, it is important to understand the firm level responses. The first two chapters of this dissertation provide a new perspective on how access to ICT shapes firms.

In addition, the rise of the digital age suggests that geographic frictions can be easily overcome nowadays. More than twenty years ago, Cairncross (1997) claimed that geographical distance would soon not play a role anymore as ICT would let any frictions disappear. However, face-to-face communication remains relevant for social and economic clustering (e.g. Gaspar and Glaeser, 1998; Storper and Venables, 2004). Even global technology companies such as Facebook and Yahoo require their employees to work in their company offices (Reses, 2013; Zuckerberg, 2016). Despite all the possibilities offered by ICT, they report substantial frictions to optimal communication and coordination if employees are not working at the same location. Indeed, Battiston et al. (2017) find that face-to-face communication remains a significant determinant of team productivity. Hence, ICT do not seem to have let "distance die". It is therefore important to understand how firms respond to the geographic frictions. Chapter 3 of this dissertation focuses on geographic frictions and their effect on the managerial organization of firms, even when ICT exist.

In chapters 1 and 2, we[1] study firm responses to getting access to one of the arguably most revolutionary information technologies: broadband internet. We provide causal empirical evidence on the impact of faster broadband internet availability on firm growth, employment, and firm skill composition. Identifying the impact of getting access to faster broadband internet empirically poses a challenge as firms self-select into the adoption of faster broadband internet. This makes it difficult to disentangle the causal effect of broadband internet speed adoption from unobserved characteristics of firms that choose to adopt faster broadband internet. To overcome this problem, we use an identification

[1]To facilitate readibility, the pronoun "we" is used throughout this thesis to refer to the author or the authors of the respective chapter.

strategy originally proposed by Falck et al. (2014). Our results show how existing firms respond to the introduction of new technologies as we compare the same firm before and after getting access to broadband internet. Thus, our results are not driven by start-ups that use broadband internet as a key component in their original business model. To the best of our knowledge, we are the first to causally identify within-firm responses to broadband internet speed availability at the time of introduction.

Our identification strategy is based on an historical accident. During the early 2000s, broadband internet was diffused over the copper wires of the established telephone network in Germany. This imposed technological restrictions: the nature of the network led to exogenous differences in the ability of establishments and firms to adopt faster broadband internet. The main distribution frames (MDFs) in the network are connected through copper wires to establishments and households. With increasing distance to the MDF, the maximum speed provided decays up to a threshold distance. Whereas Falck et al. (2014) relate the distance between the geographical middle point of a municipality to the threshold to instrument the share of households with broadband internet access, we extend their strategy by calculating and using the individual distance of each geocoded establishment to the MDF. Our set-up allows identifying establishment responses to the *first generation of broadband internet* access and *later speed upgrades* in Germany.

Chapters 1 and 2 study the responses of single-establishment and multi-establishment firms separately. As shown in chapter 1, single-establishment firms include many young firms with high growth potential. For example, they profit from decreasing costs of market access as the possibility to set up an e-business offers a cheap way to enter new markets. Understanding how these firms respond to ICT is important to draft policy that stimulates future growth. In chapter 2, we study multi-establishments. They are a very important subgroup of firms as they employ disproportionate shares of the workforce (30% in Germany; 80% in the US, see Bernard and Jensen, 2007). Their reaction to ICT thus has a big impact on aggregate labor demand.

We assemble unique datasets for our analysis. We use German social security records that provide information on establishments and employees. We combine the administrative data with firm-level information. For single-establishment firms, we use information from an establishment survey. For multi-establishment firms, we add information provided by Bureau van Dijk via a record linkage procedure. We can thus identify which establishments belong to the same firm. To implement our identification strategy, we geocode the establishments and add information on the telephone network to assess which broadband internet speed is available to each establishment.

Chapter 1 studies how access to faster broadband internet affects single-establishment firm growth. We find that firms with access to the first generation of broadband internet reduce their employment while keeping their output constant. This finding suggests increasing efficiency through broadband internet acesss. Moreover, firms with access to the first generation of broadband internet reduce the share of low-skilled employment in their workforce. Besides, we find that firms that receive later speed upgrades grow more in revenues and employment than firms that do not get access to the upgrades. This finding points towards an increasing market size effect of the access to faster broadband internet. In addition, firms with access to the later speed upgrades increase the share of medium-skilled while decreasing the share of high-skilled employment.

Our results suggest both positive efficiency and market size effects of the access to faster broadband internet. The first generation of broadband internet seems to decrease the required labor input per unit of output. The later speed upgrades seem to provide the opportunity to increase revenues and hence to expand production. In addition, our findings provide evidence for complementarity of medium-skilled labor to broadband internet.

We contribute to the literature by providing causal evidence on the effect of the access to faster broadband internet on firm growth. Our differences-in-differences approach allows ruling out time-constant firm characteristics that may otherwise explain differences between firms that do or do not get access to the new technologies. Exploiting the exogenous differences in broadband internet speed availability allows excluding any further bias of our estimates through omitted variables. We also provide supportive evidence of skill-complementarity of broadband internet as found by Akerman et al. (2015). Firms adjust their skill composition by increasing the share of more highly skilled employment as a response to faster broadband internet availability. Our findings provide a more comprehensive picture on the firm adjustments in response to getting broadband internet access.

Preface

Chapter 2 is based on joint work with Anna Gumpert, Eduardo Morales, Ezra Oberfield, and Manfred Antoni. We focus on the analysis of the impact of broadband internet speed availability on firm employment and skill composition to multi-establishment firms. Understanding multi-establishment firms may further our understanding of the impact of ICT on multinational firms: Baldwin (2016) predicts that the reduction of face-to-face costs through advancing ICT will lead to another wave of unbundling of tasks between developed and developing countries. However, to date, very little empirical evidence in this respect exists.

It is important to consider the access to faster broadband internet at the establishment level as opposed to the firm level, as adoption may vary between establishments. Firms only adopt certain broadband internet speed in those establishments where it is efficient. The impact of faster broadband internet on multi-establishment firms may vary depending on where it is adopted. For example, if the headquarters of a firm adopts faster broadband internet, this may facilitate managing subordinate establishments by reducing coordination costs and thus have repercussions on the subordinate establishment. Our identification strategy, as applied in chapter 1, allows differentiating between the impact of an establishment's own access to faster broadband internet and the respective headquarters' access.

We find that subordinate establishments respond to their headquarters' access to faster broadband internet. Subordinate establishments that already employ high-skilled labor respond more strongly to faster internet availability at their own location and at their headquarters'. In addition, subordinate establishments shift their skill composition towards more skilled labor when their headquarters gets access to faster broadband internet. Thus, our results point towards interdependence of subordinate establishments and headquarters as well as skill-complementarity of faster broadband internet.

Similar to chapter 1, chapter 2 contributes to the literature on the impact of broadband internet on firm behavior. We provide a more comprehensive picture on the skill-complementarity of broadband internet by showing that firm responses to the access to faster broadband internet depend on their initial skill composition. By showing that subordinate establishments increase the share of more skilled labor, we further provide evidence on firm adjustments in the workforce. In addition, chapter 2 extends the literature on multi-establishment firms by providing empirical evidence on the interdependence of the establishments within a multi-establishment firm. This finding suggests that establishments of multi-establishment firms should not be treated as independent units.

Consistent with the ongoing discussion on the importance of face-to-face communication, our findings in chapter 2 suggest that geographic frictions between establishments are more important than previously thought. Establishments may be interdependent because access to faster broadband internet in the headquarters may facilitate the coordination of subordinate establishments and thus have repercussions on the optimal organization of the firm. As our analysis in chapter 2 does not allow isolating the impact of easier access to information from the reduction of internal frictions within the firm, we study this second channel in a separate analysis.

In chapter 3, which is based on joint work with Anna Gumpert and Manfred Antoni, we study how geographic frictions affect firms' managerial organization. We develop a model to show that geographic frictions between the headquarters and *one* subordinate establishment affect the organization of *all* establishments of a multi-establishment firm. We assume that the CEO of a firm is a resource of limited supply that is shared among the establishments. Geographic frictions increase the costs of accessing the CEO. Hiring middle managers at an establishment releases CEO time that is reallocated across all establishments. This increases the production efficiency of the establishments and thus affects their optimal organization.

We provide empirical evidence supporting our model implications. The model explains cross-sectional differences between single and multi-establishment firm organization that we uncover using administrative data from Germany. We exploit the opening of high speed train routes to show that not only establishments directly affected by faster travel times, but also the other establishments of the firm adjust their organization. Our findings imply that empirical analyses at the establishment level may underestimate the impact of local conditions on multi-establishment firms.

This insight is particularly relevant for the literature on multi-establishment and multinational firms. Recent papers uncover that distance to the headquarters and other geographic frictions decrease investment, productivity and longevity of subordinate establishments of multi-establishment firms (e.g. Giroud, 2013; Kalnins and Lafontaine, 2013). In this literature, standard models implicitly assume that firms copy their existing operations when opening foreign affiliates (e.g., Helpman et al., 2004; Antràs and Yeaple, 2014, for a survey). To the best of our knowledge, Charnoz et al. (2015) is the only study of the impact of geographic frictions on firm organization. We provide a novel and nuanced interpretation of the regression results in light of our theoretical model.

Besides, our study contributes to the literature on firm organization and management by showing that geographic frictions are a determinant of firm organization. Previous literature focuses on firm size as the major determinant of firm organization (for an overview, see Garicano and Rossi-Hansberg, 2015).

Overall, this dissertation contributes to the understanding of the impact of ICT on firm behavior. Firms with access to broadband internet become more efficient and increase the share of more skilled employment. It further points out how, notwithstanding the powerful technologies, geographic frictions remain key determinants of firm organization.

Chapter 1

Firm Responses to High-Speed Internet

1.1 Introduction

Information and communication technologies (ICT) are general purpose technologies that enable firms to reshape their business (e.g. Brynjolfsson and Saunders, 2010). Broadband internet, in particular, is said to have revolutionized many business processes. Firms may set up e-businesses and hence increase the size of the market they can serve. Further, broadband internet facilitates file-sharing and offers new communication tools like videoconferencing. Still, causal evidence on firm growth affected by broadband internet is limited as identifying the causal impact is difficult.

This chapter studies how access to the broadband internet affects firm growth. Our set-up allows identifying firm responses to the first generation of broadband internet access and later speed upgrades in Germany. We analyze within-firm growth and workforce adjustments caused by the access to the broadband internet and the later speed upgrades using detailed social-security data.

We find that firms with access to the first generation of broadband internet reduce their employment while keeping their output constant. They specifically reduce the share of low-skilled employment in their workforce. Further, we find that firms that get access

This chapter uses the same identification strategy as chapter 2. To make each chapter self-contained, we explain the identification strategy in both chapters. I would like to thank Manfred Antoni and Florian Zimmermann at the Institute for Employment Research in Nuremberg for their great support.

to later speed upgrades grow more in revenues and employment than firms that do not get access to these upgrades. When getting access to higher internet speed, firms increase the share of medium-skilled while decreasing the share of high-skilled employment.

Theoretically, one would expect two distinct effects of broadband internet adoption on firm growth in revenues and employment. First, broadband internet is said to increase efficiency in production processes by reducing communication and coordination costs with customers and suppliers. As a result, required labor per unit of output decreases. Empirically, we test this hypothesis of increasing efficiency by regressing revenues per employee on broadband internet access.

Second, broadband internet potentially increases the size of the market a firm can serve by reducing search costs and offering new sales opportunities. Browsing the web provides a cheap way of searching for information on new markets. Also, firms can set up e-businesses that create new sales channels. As a result of the increase in market size, firms would grow in revenues and expand production. Thus, firms require more of each production factor, including labor. Empirically, we test this hypothesis by using revenues and value-added as the outcome variables to measure output. Further, we use different employment measures as outcome variables. In practice, both effects may occur simultaneously. In our analysis, we observe the net effect of potential expansions in output and employment as well as a decreasing ratio of required labor per unit of output.

Apart from its believed positive impact on growth, the expansion of broadband internet is said to contribute to a rising skill-biased technological change (Akerman et al., 2015). Policymakers should not only consider the overall employment effects of new technologies but also take the distribution effects into account. To contribute to this discussion, we study the changes in the skill composition of firms getting access to the broadband internet and later speed upgrades.

Identifying the impact of getting access to the broadband internet empirically poses a challenge as firms self-select into broadband internet adoption. This makes it difficult to disentangle the causal effect of broadband internet adoption from unobserved characteristics of firms that choose to adopt broadband internet. To overcome this problem, we use an identification strategy originally proposed by Falck et al. (2014) to compare firms with differential access to faster broadband internet before and after its introduction and later speed upgrades. During the early 2000s, broadband internet was diffused over the copper wires of the established telephone network in Germany. This imposed technological restrictions: the nature of the network exogenously led to differential access to higher broadband internet speed levels (here: digital subscriber line, DSL). The main distribution frames (MDFs) in the network are connected through copper wires to firms and households, also called the "last mile" of the network. With increasing distance to the MDF, the maximum speed provided decays up to a threshold distance. Whereas Falck et al. (2014) relate the distance between the geographical middle point of a municipality to the threshold to instrument the share of households with broadband internet access, we extend their strategy by calculating and using the individual distance of each firm to the MDF.

We exploit the fact that distance to the MDF affected broadband internet speed availability in Germany, a technical feature that firms could not anticipate. We compare firms before and after the introduction of the first generation of broadband internet by restricting our sample to firms within small bounds around the threshold distance. Moreover, our identification strategy allows studying the impact of speed upgrading in addition to the effect of access to the first generation of broadband internet. Within the group of firms that got access to the first generation of broadband internet, only a subgroup also got access to the later speed upgrades. Again, the distance to the MDF determines how fast the internet a firm receives would be. Further, our strategy allows analyzing within-firm responses to the available speed upgrades by comparing the performance of a firm before and after the introduction of the upgrades. Hence, our findings are not driven by start-ups that particularly use the broadband internet in their business model. In our specification, we control for time- and firm-fixed effects to rule out time-constant firm characteristics and common time shocks. We analyze the impact on existing firms that get access to new technologies.

For our analysis, we assemble a novel dataset: using geographic information system software[1], we geocode single-establishment firms in a survey provided by the German Employment Agency (see Heining et al., 2016). We merge the geocoded telephone net-

[1] We use QGIS version 2.18.3 (QGIS Development Team, 2017).

work (Bundesnetzagentur, 2017) to calculate the distance from each establishment to its MDF. We further combine the establishment-level data with employee-level social-security data which allows us to study changes in the workforce composition.

Looking at firms getting access to the first generation of broadband internet, we find evidence that is consistent with increasing efficiency. Firms that get access to the first generation grow less in employment than firms that do not get access. They do not grow differently in revenues leading to increasing revenues per employee. Through the lense of potential theoretical impacts, this suggests increasing efficiency, but no market size effect. We also find that firms reduce the share of low-skilled employees in their workforce pointing to skill complementarity of broadband internet. The first generation of broadband internet facilitated the exchange of e-mails with large attachments (500kB and more). Hence, firms that got access probably engaged more in digitization processes that require less low-skilled labor than previous administrative work. Separating manufacturing and non-manufacturing firms reveals that non-manufacturing firms benefit most from getting access to the broadband internet.

A subgroup of firms that got access to the first generation of the broadband internet also experienced later speed upgrades. We find positive growth effects in revenues and employment indicating a market size effect. We find no significant evidence suggesting increasing efficiency, even though the results suggest a positive effect. Firms increase the share of medium-skilled employment. In contrast to the conventional understanding of skill complementarity to new technologies, firms reduce the share of high-skilled employees. The later speed upgrades of the broadband internet provided the possibility to set up online businesses. To serve a greater market, firms probably needed to expand production. If this expansion required hiring over-proportionally more medium-skilled labor, the share of high-skilled employees would fall without a reduction in the actual number of high-skilled employees.

To the best of our knowledge, we are the first to causally identify within-firm responses to broadband internet speed availability at the time of the introduction. We contribute to the literature in three ways. First, we build on and extend the literature on growth effects of broadband internet surveyed by Bertschek et al. (2015).[2] At the firm-level, the

[2]At the country-level, the literature mostly finds positive growth effects of broadband adoption (e.g.

literature finds mixed results. Most papers studying the impact of the first generation of broadband internet find no significant effects on growth and differential effects on employment.[3]

One exception is Akerman et al. (2015) who find positive output elasticities as well as skill-biased employment and wage effects. They exploit the quasi-exogenous time variation in broadband infrastructure expansion in Norwegian municipalities due to the limited funding of a government initiative. In line with their results, we find increases in skill-biased demand for labor of firms with access to the broadband internet. This chapter adds to their findings in a number of ways. We are able to look at within-firm responses, which allow me to control for any time-constant firm characteristics. Further, we exploit exogenously given technological restrictions which allow me to compare similar firms around the threshold distance within the same municipality. Our strategy additionally allows comparing firms during the introduction of the first generation of broadband internet as well as later speed upgrades. Hence, our results provide further evidence on the economic implications of speed upgrades. Moreover, we show that broadband internet affected firms in different sectors heterogeneously. Thus, this chapter contributes to the understanding of the actual sources of growth stemming from investments in broadband infrastructure.

In another relevant paper, Canzian et al. (2015) study the impact of the second generation of broadband internet (called ADSL2+) on firm growth in rural areas in the province of Trento (Italy). They exploit a government program upgrading rural areas to higher internet speed using longitudinal firm-level data. They find large positive effects on revenues and value-added and no effect on employment. This chapter adds to their findings by including firms in both rural and urban areas and hence estimating an average effect which is of policy interest. Further, they analyze a later time period when the second major generation of broadband internet was already widespread in

Czernich et al., 2011). Further, broadband availability at the county- and zip code-level is found to have a positive effect on employment in the USA (e.g. Kandilov and Renkow, 2010; Kolko, 2012), which mostly benefits skilled workers (Atasoy, 2013).

[3] De Stefano et al. (2014) exploit a fuzzy regression discontinuity design and find no effect of the broadband internet on firm growth. Stockinger (2017) employs the instrumental variable approach developed by Falck et al. (2014) directly to study employment growth of German establishments. He finds negative effects on employment growth of establishments in manufacturing and positive effects in knowledge-intensive industries. Our rich dataset allows studying firm responses to broadband internet in more detail.

many areas. Hence, their large effects may be driven by a catch-up effect. Our analysis focuses on the time of the first introduction of the broadband internet and later speed upgrades, including the second major generation of broadband internet. Analyzing a longer time period further allows us to observe effects that only show up with a time lag. Overall, this chapter provides a more comprehensive picture of firm responses to broadband internet.

Second, we complement the literature on growth effects of ICT in general.[4] As a whole, the literature finds that productivity effects of ICT alone are very low. To fully exploit the potential of the new technologies, firms need to provide complementary factors like organizational adjustments (see e.g. Brynjolfsson and Hitt, 2000). We contribute to this literature by showing that firms that get access to the broadband internet and later speed upgrades increase the share of skilled labor.

Third, we extend the instrumental variable strategy by Falck et al. (2014) who study voting behavior. They exploit the distance between the geographical middle point of a municipality to its MDF as an instrument for the share of households in a municipality with broadband access. We build on this idea in our differences-in-differences approach by calculating the distance between each individual firm and the dedicated MDF. Hence, we approximate the access to the broadband internet at the firm-level instead of the municipality-level. Our strategy allows measuring the technological restriction more precisely and calculating the intention-to-treat effect on similar firms within the same municipality.

The chapter is structured as follows. The next section describes our empirical strategy. We explain the set-up, our empirical model, and the data we use. In section 1.3, we describe our main findings, report results from a number of robustness checks, and discuss our results. Section 1.4 concludes.

1.2 Empirical Strategy

Identifying the impact of broadband internet adoption on firm growth poses a challenge to the empirical researcher. Ideally, one would randomize technology adoption. In prac-

[4] For a detailed survey of the literature on the impact of ICT on productivity we refer to Draca et al. (2006) and Cardona et al. (2013).

tice, new technologies are usually available to everyone at the same time, and firms choose whether to adopt these. Hence, measured firm responses to new technologies would be biased due to omitted variables. Firms that select into technology adoption may simultaneously be subject to different changes correlated with firm growth, e.g., innovation activity. As a result, the impact of the new technology on firm growth cannot be identified. To overcome this problem, we exploit a technological restriction to broadband internet adoption that is orthogonal to firm characteristics. This exogenous factor led to differential access of firms to the new technology that would otherwise have selected into adopting it. We restrict our sample to similar firms that are located below and above the threshold distance to the MDF and hence may or may not receive broadband internet.

1.2.1 Set-up

Broadband expansion. The first generation of broadband internet, asymmetric digital subscriber line (ADSL), was first presented in Germany in 1999. Providing a downstream speed of 768 kBit/s and an upstream speed of 128 kBit/s, it was considered a major improvement compared to previous dial-up technologies.[5] In 2000, about 600,000 customers subscribed to the new technology. With 768 kBit/s, one could e.g. send large attachments (500kB) in an e-mail.

In later years, the technology was further improved leading to speed upgrades (see figure A.3 in the appendix), mostly in downstream. From September 2002 on, the first generation of broadband internet allowed up to 1,536 kBit/s in download speed. In April 2004, the maximum provided speed increased to 3 MBit/s in download speed and 384 kBit/s in upstream by upgrading the technology to ADSL2, i.e. the second upgrade. With 3 MBit/s, small video conferences and online meeting presentations were possible. From 2005 on, with the third upgrade, up to 6 MBit/s in downstream and 512 kBit/s in upstream were possible. With 6 MBit/s, third-party hosted applications like e-mail and data back-up could be used.

The fourth upgrade to ADSL2+ in 2006, called the second major generation of broadband internet, provided a major improvement to broadband internet provision as the

[5]For a detailed description of the technological background see Schnabel (2015).

maximum speed increased to 25 MBit/s[6] in downstream and 1,024 kBit/s in upstream. Most telecommunications providers offered up to 16 MBit/s. With 16 MBit/s, multipoint videoconferencing, remote server access, and voice over internet protocol applications came up (Columbia Telecommunications Corporation, 2010). Overall, these speed upgrades significantly improved the use of any application, especially file-sharing.

From 2006 on, another new internet technology called Very High Speed Digital Subscriber Line (VDSL) was introduced. This technology provided even higher internet speeds of up to 100 MBit/s. However, VDSL required large infrastructure investments as fiber wires needed to be installed. It therefore took several years to introduce VDSL in major cities in Germany and is still ongoing in 2018.[7]

At the time of the introduction of ADSL2+, more than 14 million customers subscribed to DSL provision.[8] Still, not every customer got access to the full potential of each technology (Bundesnetzagentur, 2011). Even in 2011, more than 22% of DSL customers received only 2 MBit/s or less. Most customers (46.3%) received between 2 and 10 MBit/s whereas only 23% received between 10 and 30 MBit/s even though ADSL2+ was already well established. Only about 8% got access to the upgrade to VDSL providing between 30 and 100 MBit/s.

Technological restrictions to broadband internet access. The early generations of broadband internet used the existing public switched telephone network (PSTN). It consisted of copper wires that could be used to transfer broadband internet. The network was constructed in West Germany in the 1960s by the state monopoly on telephone networks at that time. They aimed at providing telephone access to the universe of households. As distance is irrelevant for the quality of telephone usage, they optimized installation and maintenance costs by serving as many customers by each MDF as possible.

For broadband internet, however, the distance between a firm and the MDF matters. As shown in figure 2.1b, there is a large decay of the technologically maximal speed

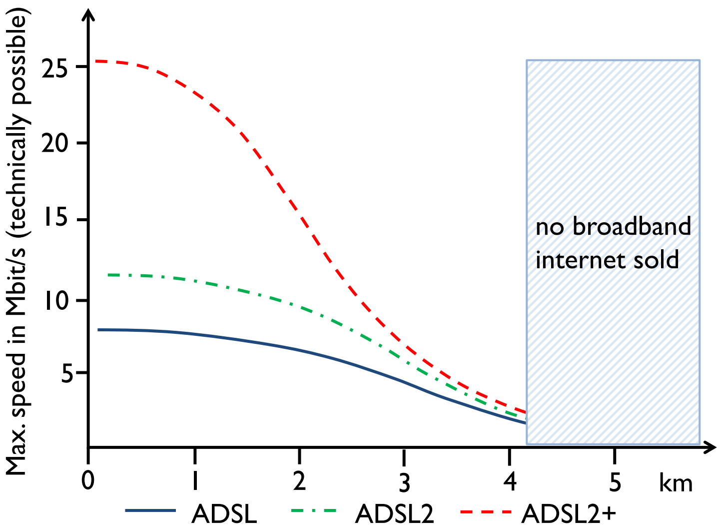
[6] Most telecommunications providers did not always provide the maximum speed (Schnabel, 2015).

[7] We provide robustness checks excluding counties where VDSL was installed until 2008 in tables A.27 and A.28 in the appendix.

[8] The number of subscriptions to DSL technologies reached 24 million in 2016. An overview of the increase in subscriptions over the considered time horizon in this chapter can be found in figure B.1.

provided with increasing length of the copper wires. For the first generation of broadband internet (ADSL), the maximum speed ranges from less than 4 to 8 MBit/s after the latest upgrade. Above the threshold distance of 4.2 km, no broadband internet was provided. Similarly, for ADSL2, the maximum speed for firms close to the MDF is highest and decays with the distance to the MDF. For ADSL2+, the speed decays even more steeply. Within the group of firms that got access to the first generation of broadband internet, only firms located between 0 and 2 km from the MDF got full access to the speed upgrade to usually 16 MBit/s. Using the network offered a cheap way to introduce the new technology. New constructions would have implied high installation costs as wires are installed subsurface in Germany.

Figure 1.1: Decay of broadband internet speed



The figure shows the decay of the maximum broadband internet speed that was technically possible with increasing distance to the MDF. The solid line shows the decay of the first generation. The line above shows the decay of the less used technology called ADSL2. The dashed line shows the decay for the second major generation of broadband internet (ADSL2+). At the time of the introduction, most telecommunications companies only provided a maximum speed of 16 MBit/s. Source: Schnabel (2015) and own illustration.

1.2.2 Empirical specification

Our basic framework consists of a fixed-effects difference-in-differences estimation equation:

$$\mathsf{y}_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{broadband\ internet\ access}_i\, D_{\mathsf{year,t}} + \alpha_i + \alpha_t + u_{it} \qquad (1.1)$$

We analyze different time-varying firm outcomes $\mathsf{y}_{i,t}$ that measure performance, size or workforce composition of firm i in year t. The dummy variable $\mathsf{broadband\ internet\ access}_i$ equals one if broadband internet provision to a firm is technically feasible and zero otherwise. It does not vary over time.

To study the impact of the access to the first generation of broadband internet, we interact the $\mathsf{broadband\ internet\ access}_i$-dummy with a post-introduction period dummy which allows us to compare firms with and without access to broadband internet before and after the introduction of the first generation of broadband internet. The post-introduction dummy is equal to one for all years from 2001.[9] To study the impact of getting access to later speed upgrades, we use dummies from year to year to trace out the timing of the effect. α_i are firm-fixed effects that control for all time-constant firm characteristics.

α_t are year-fixed effects that control for any economy-wide year characteristics. $\beta_{1,t}$ is thus an estimate for the difference between firms with and without access to broadband in each year t, holding all time-persistent firm characteristics constant and controlling for all year-fixed effects. u_{it} is the error term which is clustered at the firm-level. We balance treatment and control groups within three-digit sector groups following Iacus et al. (2009).

This equation would still suffer from omitted variable bias if we directly observed broadband internet adoption and simply included a dummy equal to one if the firm adopted broadband internet and zero otherwise. In this case, there could be unobserved time-varying factors that differ between firms that get access to the broadband internet

[9] As we observe firms as of June 30, only a very small subgroup of firms was connected to a MDF with broadband internet access on June 30, 2000. We run a robustness check in which we drop all firms in areas where MDFs were upgraded before June 30, 2000. Further, we run a robustness check in which we define the post-dummy for all years from 2000 on. We report the results for both robustness checks in table A.3 in the appendix.

and firms that do not get access. This problem can be solved by exploiting the technological peculiarities of the network as explained in sub-section 1.2.1. These technological restrictions affect the availability of broadband internet to each individual firm, and allow me to compare firms that lie just below a cutoff for broadband availability to firms that lie just above it. As we do not observe adoption of broadband internet, we calculate an intention-to-treat effect.

1.2.3 Data and descriptive statistics

For our analysis, we assemble a unique longitudinal firm-level dataset. We combine linked employer-employee data from Germany with geocoded information on the included firms and the telephone network. Germany provides an ideal setting for this study, as German social security data are rated very highly regarding availability and reliability (see Card et al., 2013).

Data. Our firm-year-level data come from an establishment survey provided by the German Employment Agency (Heining et al., 2016). This dataset reports detailed establishment information on 30 June on an annual basis from 1993 to 2014. Firms report revenues, value-added, and employment, among other topics. To follow the firms before and after the introduction of the new technologies, we look at the unbalanced panel samples from 1996 to 2005 and from 2000 to 2011. The panel samples are constructed by the German Employment Agency. A subset of the establishments surveyed every year are followed over several years. We only use single-establishment firms to exclude potential interdependencies between establishments with and without broadband internet in multi-establishment firms. We keep observations with non-missing sales information reducing our dataset by about 20%.

Further, we merge linked employer-employee data from social security records. The individual-level data come from the Sample of Integrated Labor Market Biographies (SIAB) (Antoni et al., 2016). They include detailed current information on the employees of the firms in the survey. For every employee, we observe individual demographic information like education, wages, occupations, and employment status. Besides, we calculate the age and tenure of each employee. We impute missings in the education variable following Fitzenberger et al. (2005). We define individuals with a university degree as highly skilled, individuals with vocational training or a university entrance qualification as medium-skilled and those without any training or university entrance qualification as low-skilled. We consider full-time employees between age 18 and 65.

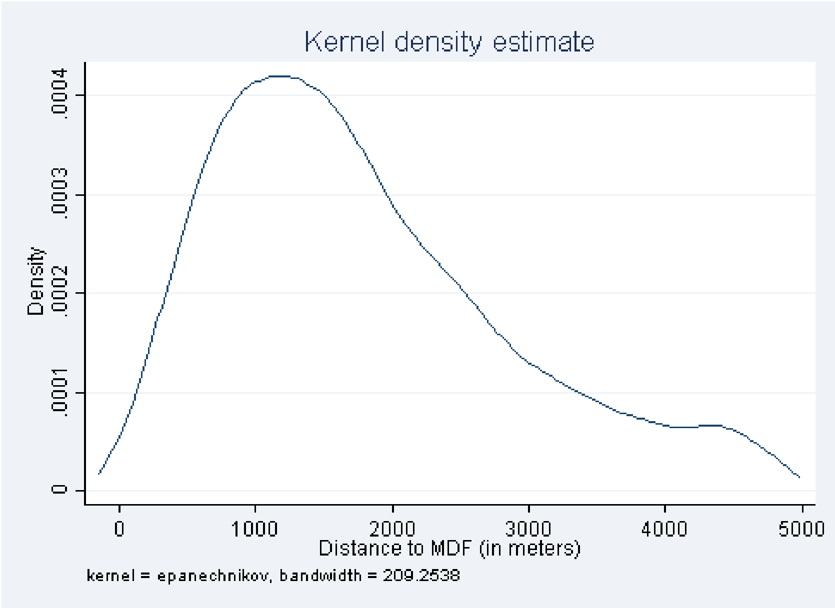
To approximate individual broadband internet access, we geocode each establishment based on the address included in the social security records. Using the geographical information on the local telephone networks provided by Bundesnetzagentur (2017), we allocate each establishment to a local network using geographic information system software. Lacking data on the actual connection of firms to MDFs, we define the closest MDF within the local network as the relevant one. To approximate the length of the copper wires, we calculate the distance via roads between the firm and the MDF using the map of road networks provided by OpenStreetMap (2017) (version as of March 2015). The distance is calculated based on the cross-sectional information provided in 2015 and does not vary over time. As the copper wires were installed subsurface, they are usually located next to roads where opening the ground is easiest.

Sample definitions. MDFs are not randomly located but rely on existing infrastructure. We argue that firms did not anticipate that distance to the MDF would matter so they did not locate strategically within small bounds. Large distances, however, might reflect very different firms. Our strategy allows excluding firms that are far away from the MDF that may grow differently after the introduction of the broadband internet because of other reasons. Figure 1.2 shows that the distribution of the distance to the MDF in the panel sample from 2000 to 2011 within 4.7 km or less is skewed to the left. 50% of firms are located within 1.4 km from the MDF, 90% lie within 3.2 km.

Our determination of broadband internet access allows us to approximate broadband internet availability for each firm but still suffers from measurement error. As we do not have exact information on the location of the copper wires, we cannot determine the exact length of the copper wires between the firms and their dedicated MDFs. To reduce potential measurement error, our control groups include a “donut”, i.e., we leave out firms that are located within a very small bound at the threshold.

Figure 1.3 summarizes our main samples. To study the impact of access to the first

Figure 1.2: Distribution of distance to MDF



The figure shows the kernel density of the distance between firms and their dedicated MDF. Own illustration.

generation of broadband internet, we compare firms below and above the threshold at 4.2 km.[10] Our treatment group consists of firms with a distance to the MDF between 2 and 3.5 km. These firms almost received the maximum speed provided by ADSL. Our control group includes firms with a distance to the MDF between 4.2 and 5.7 km. These firms are arguably comparable as they are located very close to the treated firms and did not anticipate the implications of locating further away. These firms did not get access to the broadband internet at all. We leave out firms between 3.5 and 4.2 km from the MDF. These firms still got access to the broadband internet but did not experience the full potential speed (as shown in figure 2.1b).[11]

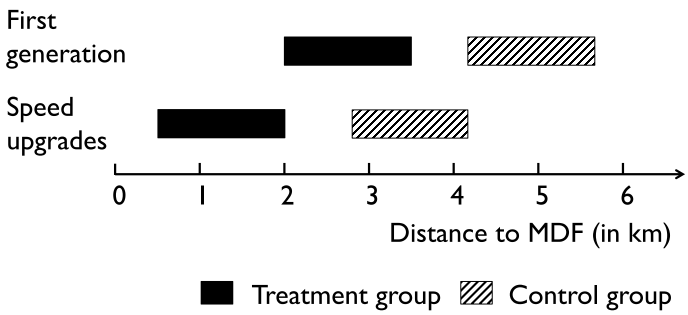
For the later speed upgrades, we exploit the fact that broadband internet speed availability differed between firms. Access to the upgrades again depended on the distance to the MDF, but the decay of the provided speed was more pronounced for shorter dis-

[10]As pointed out by Falck et al. (2014), there are other factors that also determine the maximum speed. Thus, we use this as a fuzzy threshold.

[11]In the appendix, we show the results of our second control group consisting of firms between 3.5 and 5 km. A subgroup of these firms received slow ADSL whereas firms located more than 4.2 km from the MDF did not get access to the broadband internet

tances. Within the group of firms that got access to the first generation of broadband internet, only a subgroup also got full access to the speed upgrades. After the upgrade to ADSL2+, for example, firms located between 0 and 2 km from the MDF usually received 16 MBit/s from 2006 on. Above 2 km, the maximum speed decays up to the 4.2 km threshold of any broadband internet provision. We compare firms below and above the threshold of 2 km. Our treatment group comprises firms with a distance between 0.5 and 2 km. Our control group consists of firms between 2.8 and 4.2 km. For these firms, the speed upgrades hardly increased their internet speed. We leave out firms located between 2 and 2.8 km from the MDF. These firms did not get access to the full potential of the speed upgrades. An illustration of the sample definition to analyze the impact of speed upgrades is shown in figure 1.3.[12] Figure A.1 in the appendix illustrates the distribution of MDFs and the respective treatment and control groups.

Figure 1.3: Sample definitions

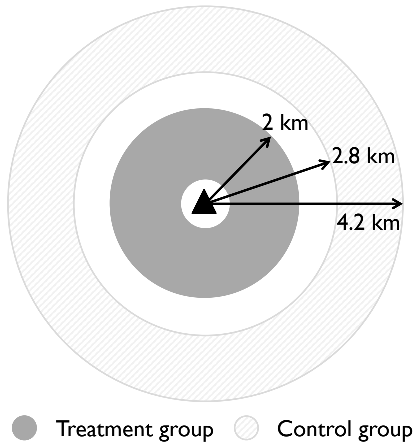


The figure shows the defined treatment and control groups for each sample by technology. The treatment group for the introduction of the broadband internet consists of firms located between 2 and 3.5 km from the MDF. The control group includes firms between 4.2 and 5.7 km. The treatment group for later speed upgrades consists of firms between 0.5 and 2 km. The control group includes firms between 2.8 and 4.2 km. Own illustration.

Descriptive statistics. For our analysis, we use different outcome variables to estimate the impact of the availability of broadband internet and later speed upgrades on firm performance, employment, and skill composition. To measure firm performance, we

[12] In sub-section A.2.7 in the appendix, we further show the results of a second control group including firms between 2 and 3.5 km. They consist of firms that are most closely located to the treatment group. These firms are very similar to the treatment group as they did not anticipate the threshold. These analyses, however, are more likely to suffer from measurement error as we do not have information on the exact location of the copper wires.

Figure 1.4: Illustration of treatment and control groups for speed upgrades



The figure illustrates the determination of treatment and control group based on the distance to the main distribution frame for speed upgrades. The black triangle resembles an exemplary main distribution frame. The circles show the areas defining the treatment and control group. White areas show the left out "donut" circles. The dark circle resembles the treatment group of firms located between 0.5 and 2 km from the MDF. The shaded circle shows the control group from 2.8 to 4.2 km around the MDF. The circles would show exemplary borders if distances were calculated via straight lines. They do not resemble the actual thresholds as the distances are calculated via roads. Own illustration.

mainly use information from the survey data. Firms report their revenues as well as the share of purchased inputs. We further calculate annual value-added generated by the firm based on this information. To approximate efficiency, we construct two measures. Based on the survey data, we calculate revenues per employee. Based on the information on the employees from social-security data, we calculate the revenues per full-time employee. We further use full-time employment and total employment as outcome variables measuring employment size. Besides that, we use information from the survey on monthly wage sums as well as social-security data to construct daily wage sums. The main difference between the two employment measures is that the survey data includes part-time employees. Using the social-security data, we further calculate the annual skill composition using the shares of three different skill groups in the full-time employment.

Table 1.1 shows the descriptive statistics for the analysis of the impact of the introduction of the broadband internet. We pool all years before the introduction in the pre-treatment period. Firms in the treatment group perform better regarding revenues and value-added. They are also larger in total employment and wage sum. The difference is economically small. Controlling for firm-fixed effects in our specification rules out any time-persistent characteristics. Hence, different initial situations of firms and treatment and control group only cause a problem if they affect growth rates which would lead to biased estimates.

Table 1.1: Descriptive statistics of outcome variables pre-treatment, pooled sample

	Treatment group			Control group			Difference	
	Mean	SD	N	Mean	SD	N	(4) - (1)	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Performance								
Log(revenues, survey)	14.53	1.89	887	14.28	1.64	343	−0.25	0.03
Log(value added, survey)	13.73	1.83	675	13.48	1.51	271	−0.25	0.04
Share of purchased inputs	0.48	0.22	684	0.44	0.21	280	−0.04	0.02
Log(revenues per FTE)	12.78	1.73	818	12.49	1.34	313	−0.29	0.01
Log(revenues per empl.)	11.32	0.84	816	11.34	0.79	312	0.02	0.71
Employment								
Log(full-time empl., SIAB)	1.68	1.49	887	1.71	1.37	343	0.03	0.74
Log(total empl., survey)	3.37	1.53	887	3.09	1.35	343	−0.28	0.00
Log(daily wage sum, SIAB)	5.88	1.61	887	5.88	1.51	343	0.00	1.00
Log(monthly wage sum, survey)	10.58	1.77	853	10.25	1.54	331	−0.33	0.01
Skill composition								
Share low-skilled	0.05	0.16	887	0.06	0.17	343	0.01	0.26
Share medium-skilled	0.83	0.26	887	0.88	0.23	343	0.05	0.01
Share high-skilled	0.10	0.22	887	0.06	0.16	343	−0.05	0.00

				Shares	
Units of observation	Treat	N	IDs	Urban	Manuf.
Firms	0	595	79	.33	.31
Firms	1	1,760	235	.47	.44
Employees	0	43,566	8,355	.34	.60
Employees	1	393,772	75,522	.55	.24

This table shows the descriptive statistics of the firms in panel sample 1996 to 2005 in the pre-treatment period (1996 to 2000). Data sources are reported in the table. The shares of skill groups are calculated based on the composition in SIAB. The p-value in column 8 indicates whether the means of the treatment and control groups are significantly different from each other.

Table 1.2 shows the descriptive statistics in the first panel year for the analysis of the speed upgrades. Comparing the treated firms located around 0.5 to 2 km from the MDF to the control group shows that these groups of firms are very similar before the speed upgrades. Table A.1 in the appendix shows the descriptive statistics for non-manufacturing firms separately.

Table 1.2: Descriptive statistics of outcome variables in 2000, pooled sample

	Treatment group			Control group			Difference	
	Mean	SD	N	Mean	SD	N	(4) - (1)	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Performance								
Log(revenues, survey)	14.41	1.97	1,046	14.45	1.90	284	0.13	0.75
Log(value added, survey)	13.57	1.98	938	13.63	1.84	259	0.15	0.66
Share of purchased inputs	0.49	0.23	953	0.48	0.22	265	−0.01	0.63
Log(revenues per FTE)	12.69	1.78	1,012	12.55	1.73	276	0.04	0.24
Log(revenues per empl.)	11.40	0.86	974	11.38	0.76	264	−0.03	0.23
Employment size								
Log(full-time empl., SIAB)	1.69	1.53	1,046	1.91	1.58	284	0.10	0.03
Log(total empl., survey)	3.20	1.66	1,046	3.28	1.55	284	0.10	0.47
Log(daily wage sum, SIAB)	5.92	1.69	1,046	6.14	1.72	284	0.12	0.05
Log(monthly wage sum, survey)	10.46	1.98	1,013	10.58	1.84	278	0.17	0.33
Skill composition								
Share low-skilled	0.07	0.19	1,046	0.06	0.16	284	−0.01	0.44
Share medium-skilled	0.83	0.27	1,046	0.82	0.24	284	−0.00	0.93
Share high-skilled	0.09	0.21	1,046	0.11	0.21	284	0.02	0.27

				Shares	
Units of observation	Treat	N	IDs	Urban	Manuf.
Firms	0	2,516	343	.42	.55
Firms	1	10,006	1,337	.54	.47
Employees	0	264,340	50,688	.48	.47
Employees	1	629,766	120,759	.64	.37

This table shows the descriptive statistics of the firms in panel sample 2000 to 2011 in the first year. Data sources are reported in the table. The shares of skill groups are calculated based on the composition in SIAB. The p-value in column 8 indicates whether the means of the treatment and control groups are significantly different from each other.

1.2.4 Discussion of identification strategy

In our identification strategy, we exploit technological restrictions to broadband internet speed adoption that firms could not anticipate or directly influence. We discuss several potential concerns and argue that our estimator would, if anything, be biased towards zero.

First, the remaining threat to our identification is that unobserved time-varying shocks may impact firms in the treatment and control differently. Hence, we check if our parallel trend assumption holds for all years before treatment. As shown in tables A.4 and A.5 in sub-section A.2.1 in the appendix, firms in the treatment and control group have similar trends before the introduction of the broadband internet in the panel sample from 1996 to 2005. The coefficients of the interaction terms of the treatment dummy and year dummies are not significantly different from zero for the years before 2001. For the speed upgrades, we show yearly coefficients in all tables. In the panel sample from 2000 to 2011, firms in the treatment and control group have similar trends before the speed upgrades starting in 2003. Hence, the parallel trend assumption seems to hold.

Second, as we approximate the length of the copper wires between the MDF and the firm, our estimates may suffer from measurement error. This problem should, however, be reduced in our "donut" samples. We exclude firms that could be falsely allocated to the treatment or control group by calculating a shorter or longer distance to the MDF within the respective range. Firms that are still included and allocated to the wrong group bias our estimates towards zero. Lacking data on the road network in the early 2000s or even 1960s, we use the OpenStreetMap version as of March 2015. Hence, we may falsely include roads that did not exist when the copper wires were actually installed. In this case, we would underestimate the length of the copper wire. Hence, we may falsely classify an untreated firm as treated which would bias our estimates downwards.

Third, one might be concerned about non-compliance in order to scale the intention-to-treat to an average treatment effect on the treated. On the one hand, some firms might adopt broadband internet even if they belong to the control group. Indeed, firms could lease individual lines but only at high costs. In 2004, firms had to pay more than 2 million USD PPP annually for a leased line to the telephone backbone to receive 2 MBit/s (OECD, 2005). Further, anecdotal evidence suggests that even large multinational firms refused to pay for individual broadband infrastructure. The Italian multinational small appliance manufacturer De Longhi moved its German subsidiary to a different city due to slow internet connection (Koehler, 2012). As a consequence, we consider the group of always-adopting firms supposedly small. If it existed, it would bias our estimates downwards.

On the other hand, some firms may not adopt broadband internet even though it is technologically possible. This group is more likely to exist in our analyses of the first generation of the broadband internet when firms had to set up an explicit contract with the telecommunications provider. For the later speed upgrades, however, firms had little incentives not to take advantage of higher speeds of the existing technology. Further, Bertschek et al. (2013) show that 98% of the firms in their sample use the internet in 2002. Hence, general technology adoption is very high. Firms that had the possibility to take up broadband internet had little incentives to stay with inferior dial-up technologies.

1.3 Results

In our analysis, we find evidence that is consistent with both our hypotheses on market size and efficiency effects of the broadband internet. First, we find that firms that get access to the first generation of broadband internet demand less labor per unit of output, especially employing less low-skilled labor. This result is consistent with increasing efficiency. Our results show that this effect is driven by non-manufacturing firms. Hence, we report our results for this sub-sample as our main findings. Our results for manufacturing firms are reported in the appendix (tables A.9 to A.11). Second, we show that non-manufacturing firms that get access to later speed upgrades grow more strongly in revenues and employment suggesting a positive market size effect. They particularly increase the share of medium-skilled employment. We show our main results in sub-section 1.3.1. Further, we challenge our results in several robustness checks reported in sub-section 1.3.2. We discuss their economic significance in sub-section 1.3.3.

1.3.1 Main findings

Figure 1.5 reports the results of the first generation of broadband internet for non-manufacturing firms. The coefficients show the results of the interaction term of the $\textsf{broadband internet access}_i$- and a $D_{\textsf{post,t}}$-dummy. The $D_{\textsf{post,t}}$-dummy is one for all years after 2000 and zero otherwise. Table A.6 in the appendix shows the results for the impact of the broadband internet availability in the pooled sample as well as for non-manufacturing and manufacturing firms separately. Comparing the results reveals that the effects are mostly driven by firms in non-manufacturing sectors. The signs of the coefficients for manufacturing firms, however, are in line with our findings even though they are not always significant.

We find that firms with access to the broadband internet have 9% higher revenues per employee than firms without access (significant on the 5%-level). The average firm generates revenues of about 80,800 Euro per employee before treatment. A firm that gets access to the broadband internet would generate revenues of more than 88,000 Euro per employee on average after the introduction ceteris paribus. We find no effect on revenues and value-added.

Further, we find that firms that get access to the broadband internet employ 7.5% less full-time labor after the introduction of the broadband internet (significant on 1%-level). A firm with about 13 employees before treatment would hence employ one employee less after the introduction than a similar firm without access to the broadband internet. Similarly, we find that their wage sum is smaller after getting access compared to the control group. In addition to the overall effects on employment, broadband internet access also affects the composition of the workforce with respect to their skills. Analyzing the shares of employees by skill group reveals that low-skilled employees are negatively affected by the availability of broadband internet. Firms that get access to the broadband internet employ a 2 percentage points smaller share of low-skilled labor.

For the analysis of the effects of the speed upgrades, we report our main findings for non-manufacturing firms in figures 1.6 and 1.8 as well as tables A.7 and A.8 in the appendix. We use model (1) directly including yearly coefficients to trace out the effect over time. Figure 1.6 shows the results of the effects of the speed upgrading on firm performance of non-manufacturing firms. We find that firms that get access to the speed upgrades grow more in revenues and value-added compared to the control group. As an example, treated firms sell 15% more in 2006 than non-treated firms. The yearly effect remains stable within a range between 13% in 2003 and 19% in 2007. Besides, we find no effect on revenues per employee. Table A.9 shows that the effects are similar for manufacturing firms.

Figure 1.8 shows the results of the effects of the speed upgrades on employment of non-manufacturing firms. We find that firms that get access to the speed upgrades grow more in employment. In 2006, they were 12% larger in total employment than firms that did not get access. They also report higher wage sums. In 2004, firms that get access to the speed upgrades paid a 20% larger wage sum. The effect drops down to

12% in 2006 and then rises to 20% in 2008 and 2009 again. The coefficients for total employment are larger and significantly different from zero whereas the coefficients for full-time employment are not significant. One potential explanation would be that firms hire more part-time employees to expand production. As reported in table A.10, we find similar effects for manufacturing firms.

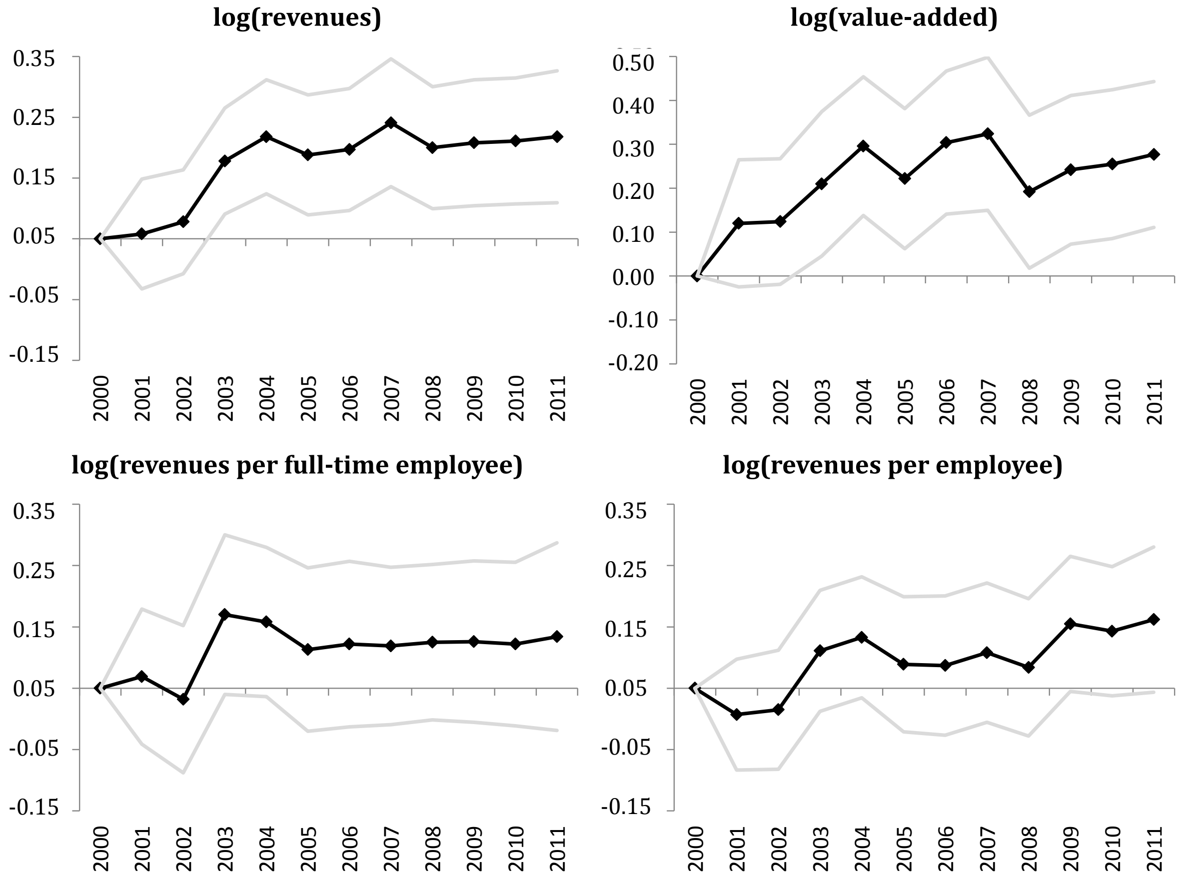
The access to later speed upgrades leads to an over-proportional increase in the use of medium-skilled labor. The share of medium-skilled employees increases, whereas the share of high-skilled employees decreases after the speed upgrades. Our results support the idea that medium-skilled labor serves as a complement to ICT as found in the literature (see, e.g. Brynjolfsson and Hitt, 2000). In contrast to the classical interpretation of skill-biased technological change, we find that the share of high-skilled employees also decreases. For manufacturing firms, we find opposite results as shown in table A.11. Firms that get access to the speed upgrades decrease the share of medium-skilled and increase the share of high-skilled employees.

To summarize, we find that firms benefit from the introduction of both technologies. The first generation of broadband internet allows firms to become more efficient employing less labor per unit of output. The later speed upgrades increase output and employment. Both the access to the first generation of the broadband internet as well as to later speed upgrades lead to an over-proportional increase in the demand for medium-skilled labor in comparison to low- and high-skilled labor in non-manufacturing firms.

Figure 1.5: Regression results of the introduction of the broadband internet

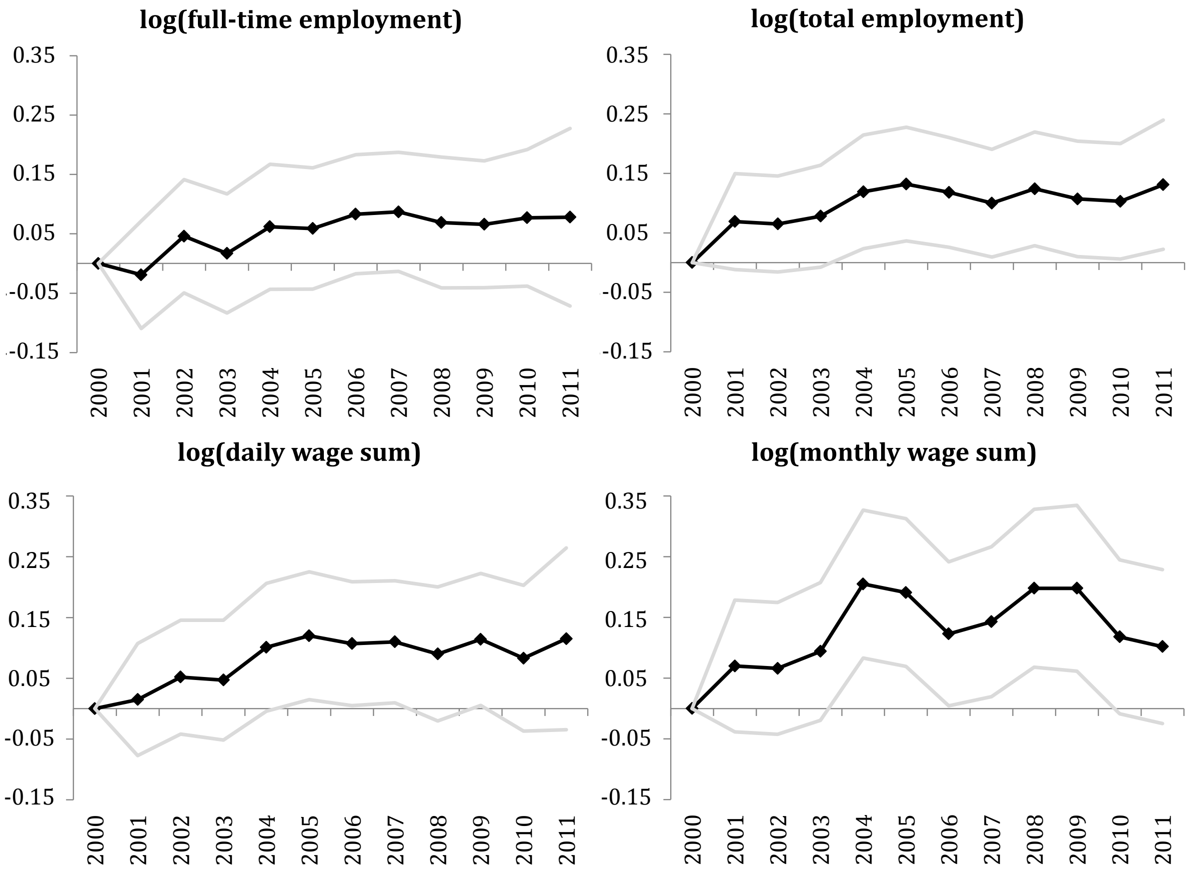
This figure shows the results from the fixed-effects difference-in-differences estimation using 1996 to 2000 as pre-treatment period and 2001 to 2005 as the treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_1 \, \textsf{broadband internet access}_i \, D_{post,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{post,t}$ is an indicator variable for the treatment period from 2001 on. $\textsf{broadband internet access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results of β_1. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, we use the weights suggested by Iacus et al. (2009). Significance level chosen at p < 0.10. The sample only contains non-manufacturing firms. Regression results reported in table A.6 in the appendix.

Figure 1.6: Regression results of speed upgrades on performance



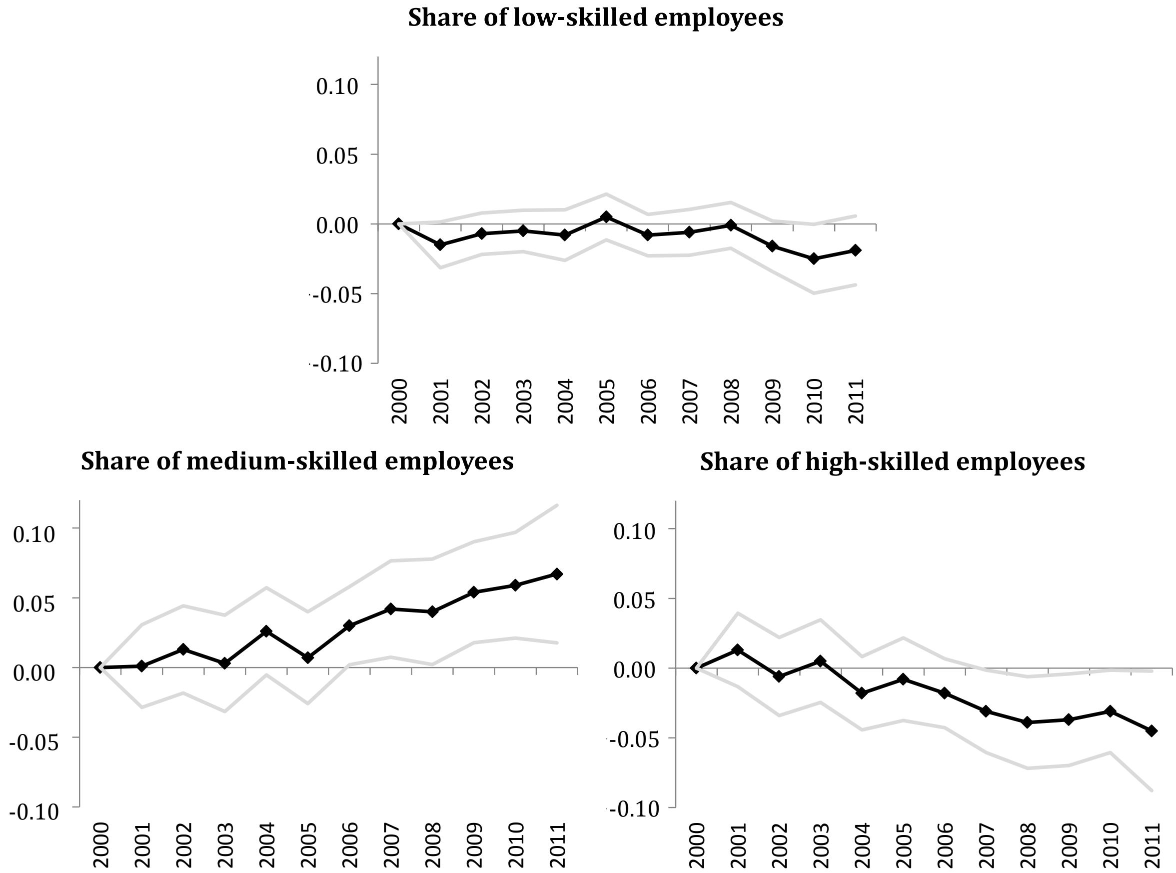
The figure shows the regression results of model (1) using the panel sample from 2000 to 2011. The upper left graph shows the results of log(revenues) as the outcome variable. The upper right graph shows the results of log(value-added) as the outcome variable. The bottom left graph shows the results of log(revenues per full-time employee) as the outcome variable. Full-time employees are counted in the administrative data. The bottom right graph shows the results of log(revenues per employee) as the outcome variable. The total number of employees is reported in the survey data. The treatment group consists of firms between 0.5 and 2 km from the MDF. The control group consists of firms between 2.8 and 4.2 km from the MDF. The samples only contain non-manufacturing firms. Grey lines mark the confidence intervals at the 10% significance level. Own illustration.

Figure 1.7: Regression results of speed upgrades on firm size



The figure shows the regression results of model (1) using the panel sample from 2000 to 2011. The upper left graph shows the results of log(full-time employees) as the outcome variable. The upper right graph shows the results of log(total employment) as the outcome variable using the data from the survey. The bottom left graph shows the results of log(daily wage sum) as the outcome variable. The bottom right graph shows the results of log(monthly wage sum) as the outcome variable using the data from the survey. The treatment group consists of firms between 0.5 and 2 km from the MDF. The control group consists of firms between 2.8 and 4.2 km from the MDF. The samples only contain non-manufacturing firms. Grey lines mark the confidence intervals at the 10% significance level. Own illustration.

Figure 1.8: Regression results of speed upgrades on the skill composition



The figure shows the regression results of model (1) using the panel sample from 2000 to 2011. The upper graph shows the results of the share of low-skilled employees (i.e. without vocational training or university entrance qualification) as the outcome variable. The bottom left graph shows the results of the share of medium-skilled employees (i.e. with vocational training or university entrance qualification) as the outcome variable. The bottom right graph shows the results of the share of high-skilled employees (i.e. with a university degree) as the outcome variable. The treatment group consists of firms between 0.5 and 2 km from the MDF. The control group consists of firms between 2.8 and 4.2 km from the MDF. The samples only contain non-manufacturing firms. Grey lines mark the confidence intervals at the 10% significance level. Own illustration.

1.3.2 Robustness checks

As discussed in sub-section 1.2.4, one might be concerned about some confounding factors biasing our results. To take these concerns into account, we run several robustness checks. Table A.12 summarizes the results of most of the following robustness checks for the introduction of the broadband internet. We report the results of the robustness checks on speed upgrading in separate tables. The results support our main findings.

First, we run the regressions using the distance to the MDF as a continuous treatment variable instead of the dummies. We report the results in tables A.13 to A.15 in the appendix. Our results confirm our previous findings. Further, we repeat the regressions on the samples without considering the "donut". Hence, we compare firms that are located within even smaller bounds from the MDF, as one may be concerned about the distance to the MDF being correlated with other time-varying firm characteristics. For the analysis of broadband internet, we use the same sample of firms as the treatment group as in the main results but firms between 3.5 and 5 km as the control group. The results are presented in table A.16 in the appendix. We find similar negative employment effects as in the "donut" sample. However, we do not find comparable positive effects on revenues per employee.

Table A.17 and table A.18 report the results of the "non-donut" samples for the analysis of later speed upgrades. We use firms between 2 and 3.5 km as the second control group. We find limited evidence on employment effects as shown in table A.17 but increasing revenues per employee as reported in table A.18. To understand these results, however, one needs to take into account that these estimations are more likely to suffer from measurement error.

Second, firms may strategically locate close to the MDF, or the MDF may be installed close to specific firms. As explained above, the network in West Germany was installed in the 1960s. Hence, MDFs could be located close to firms that existed at that time. As a robustness check, we run separate regressions on firms founded before and after the 1990s when first internet technologies were introduced. Table A.19 shows the results of the speed upgrades for the firm size of firms founded in 1992 or later. Firms that get access to the speed upgrades significantly increase the wage sum. Table A.20 shows that results of increasing revenues of treated firms are robust. Table A.21 shows the results on the firm size of firms founded in 1992 or before. Indeed, firms in the treatment group already grow faster before treatment. However, results on firm performance in table A.22 show that the effect on revenues does not kick in earlier than expected. For the first generation of broadband internet, we find similar results as in the main specification. However, the results of changes in employment are not significant. This result may also be caused by the decrease in sample size.

Further, we run separate regressions excluding East Germany where the MDFs were installed only in the 1990s. Table A.23 shows the results of firm size and table A.24 shows the results of the performance measures. We find our results to be weaker for firm size but similar for performance. For the introduction of the broadband internet, the sample size decreases to around 200 observations. Hence, the results need to be interpreted with caution: treated firms become smaller in revenues and value-added than non-treated firms.

Third, we check for potential measurement error because of lagged technology upgrading. If the MDF did not provide broadband internet, e.g. because it was not upgraded to provide the new technology yet, firms around the MDF would falsely be classified as receiving broadband internet. Hence, we use information on the share of households with broadband internet from 2005 to 2008 provided by the Federal Ministry for Economic Affairs and Energy (Federal Ministry of Economics and Technology, 2009). As a robustness check, we exclude municipalities with very low shares of households with broadband internet access in 2005. Again, we find our results to be robust to this restriction as shown in table A.25 for the results on firm size and in table A.26 for the performance measures. For the introduction of the broadband internet, we find similar results. The coefficients are not always significantly different from zero which may also be due to the small sample size. Further, our estimates may be biased by the introduction of VDSL. We therefore exclude all counties in which VDSL was introduced until 2008. Tables A.27 and A.28 in the appendix show the results supporting our main findings.

Besides, we run additional regressions with different outcome variables to check for potential mechanisms how firms adjust to getting access to the broadband internet and later speed upgrades. In Appendix A.2.9, we report the results for the regressions using exit (table A.29) and dummy equal to one if a firm invests in ICT (table A.30 for the panel from 1996 to 2005 and table A.31 for the panel from 2000 to 2011) as outcome variables. We find no effect on these variables.

1.3.3 Discussion of results

Our results are generally in line with the findings of the two most related papers. Similar to the analysis by Canzian et al. (2015) for a region in Italy, we analyze the impact of the availability of faster broadband internet on firm growth in revenues and value-added. Whereas Akerman et al. (2015) show skill-biased labor market implications and output elasticities caused by broadband internet, we complement their paper by studying the firm-level adjustments to the skill composition of the workforce. Also, we offer several detailed insights providing a more comprehensive picture of the impact of broadband internet on firm growth and employment.

In line with Canzian et al. (2015), we find large positive effects of the access to the speed upgrades like ADSL2+ on revenues and value-added. For our sample of non-manufacturing firms, we find that firms with faster internet generate about 15% more revenues than firms in the control group. Canzian et al. (2015) find that revenues increase by 40%. This large result may be driven by the fact that they study the impact of ADSL2+ at a time when many firms in other regions already had access to the new technology. Hence, firms in their sample may be experiencing a catch-up effect as the new technology was already established in the market. Our coefficients for value-added are comparable to the results in Canzian et al. (2015). However, they do not find an effect on employment whereas we find large positive effects on employment. Firms that get access to the speed upgrades employ more than 10% more labor than the control group.

Similar to Akerman et al. (2015), we find evidence for skill complementarity of broadband internet. We find that firms reduce the share of low-skilled employment when they get access to faster internet. Akerman et al. (2015), however, do not find positive labor market effects for medium-skilled labor whereas we find that firms over-proportionally demand more medium-skilled labor. One needs to note, however, that our definitions for medium-skilled labor differ. They define individuals with a high-school but no college degree as medium-skilled. As defined in sub-section 1.2.3, we define individuals with vocational training but without a university degree as medium-skilled. In contrast to the classical interpretation of skill complementarity, we also find that firms reduce the share of high-skilled labor when they get access to the speed upgrades. Our results are consistent with an interpretation of individuals with vocational training being the actual complement to broadband internet. Akerman et al. (2015) find positive and significant labor market effects and output elasticities for individuals with at least a college degree. In Germany, however, many occupations that require a college degree in other countries are taught in vocational training. Further, we estimate the effect of broadband internet on existing firms that get access to the broadband internet or later speed upgrades. Their result may be driven by new firms that enter markets where broadband internet is installed and employ a large share of high-skilled labor. Both analyses contribute to the understanding of the overall effect of broadband internet on firm growth.

Our results on the timing of the impact of the speed upgrades reveal that the effect kicks in earlier than the large upgrade to ADSL2+. This finding suggests that treated firms already benefited from earlier smaller speed upgrades. Firms that are closer to the MDF also received faster internet speed before the upgrade to ADSL2+. One might be concerned about the validity of the comparison of treatment and control group if these groups are already different before treatment. As shown in table 1.2 discussed above, however, firms in the treatment and control group were very similar before the introduction of the broadband internet in 2000. Hence, these differences in the years 2003 to 2005 are probably driven by the treatment regarding speed upgrades.

This finding further raises the question on which speed upgrades matter to stimulate firm growth. Policymakers should take into account whether firms react similarly to upgrades to 6 or 16 MBit/s. Considering the results in the sample comparing firms located between 0.5 and 2 km to firms located between 2 and 3.5 km (see sub-section A.2.7 in the appendix) one might conclude that already smaller speed upgrades had a large impact on firm growth. In this sample, firms in the control group also receive between 6 and 16 MBit/s compared treated firms receiving 16 MBit/s. Hence, the difference in internet speed between treatment and control group is very small. If such a small difference in speed is not decisive, this may explain why the results of this sample provide very limited evidence on firm growth caused by the speed upgrades.

1.4 Conclusion

New technologies like ICT are believed to drive future economic growth. As policymakers expect external effects from investments in ICT, technology infrastructure is partly publicly financed in many countries. The German government, for example, decided to spend 100 billion Euro on expanding broadband infrastructure from 2017 to 2025 (as stated by the Federal Ministry of Transport and Digital Infrastructure (Federal Ministry of Transport and Digital Infrastructure), 2017). Firms are expected to reshape their business adapting to an era of digitalization. So far, causal evidence of the impact of broadband internet on firm growth is limited.

This chapter studies the impact of the first generations broadband internet and later speed upgrades on firm growth and employment. In particular, we analyze the effect of getting access to the first generation of broadband internet and latter speed upgrades. We exploit a natural experiment of technological restrictions which implied that not all firms had access to the broadband internet. We study within-firm growth in output, employment, and adjustments to the skill composition of the workforce. Our results suggest that firms benefit from increasing internet speed. Upgrading the internet speed leads to firm growth in revenues and value-added and increases employment. Our results confirm the findings that broadband internet is a skill-biased technology found by previous literature. However, we find very limited evidence of substitution of low-skilled employees but rather an increase in medium-skilled employment of growing firms.

Chapter 2

The Impact of Broadband Internet Availability on Multi-Establishment Firms

2.1 Introduction

Information and communication technologies (ICT) are general purpose technologies that enable firms to reshape their business (e.g. Brynjolfsson and Saunders, 2010). Broadband internet, in particular, is believed to have revolutionized many business processes. It provides firms the opportunity to set up e-businesses and hence to increase the size of the market they serve. Further, broadband internet facilitates file-sharing and offers new communication tools like videoconferencing. The impact of broadband internet on multi-establishment firms is particularly interesting for policy makers due to their large employment share.[1] Hence, it is important to understand the impact of broadband internet on employment and skill composition of these firms. However, we know little about the impact of broadband internet on multi-establishment firms.

Identifying the impact of ICT, like broadband internet, on firms poses a challenge

This chapter is based on joint work with Anna Gumpert, Eduardo Morales, Ezra Oberfield, and Manfred Antoni. It uses the same identification strategy as chapter 1. To make each chapter self-contained, we explain the identification strategy in both chapters.

[1]In our data, they represent only a small share of all firms (about 9%) but employ more than one third of all employees subject to social security.

as firms self-select into technology adoption. This makes it difficult to disentangle the causal effect of the adoption of a new technology from factors that are observable to the firm but not to the researcher. Identification is even more challenging in multi-establishment firms, because adoption may vary at the establishment-level. Firms only adopt higher broadband internet speed in those establishments where it is efficient. The impact of faster broadband internet on multi-establishment firms may vary depending on where it is adopted. For example, if the headquarters (HQ) of a firm adopts faster broadband internet, this may facilitate managing subordinate establishments by reducing coordination costs and thus have repercussions on the subordinate establishment.

In this chapter, we study the impact of the availability of broadband internet at different speed-levels on the employment and the skill composition of German multi-establishment firms. We exploit technological peculiarities of the broadband infrastructure that cause differences in the availability of broadband internet speed levels at the establishment-level. Our strategy allows differentiating between the impact of an establishment's own access to faster broadband internet and the respective HQ's access. We find that subordinate establishments grow faster with their HQ's access to faster broadband internet. Further, we find that subordinate establishments that already employ high-skilled labor respond more strongly to faster internet availability at their own location and at their HQ's. Besides, they shift their skill composition towards more skilled labor when their HQ gets access to faster internet. Our results suggest that establishments in multi-establishment firms are interdependent. We further provide supportive evidence for skill-complementarity of faster broadband internet.

We exploit an identification strategy proposed by Falck et al. (2014). During the early 2000s, broadband internet was diffused over the copper wires of the established telephone network in Germany. This imposed technological restrictions: the nature of the network exogenously restricted broadband availability for some establishments. The main distribution frames (MDFs) in the network are connected through copper wires to establishments and households. With increasing distance to the MDF, the maximum speed provided decays up to a threshold distance, a technical feature that firms could not anticipate. Distance to the MDF affects both the availability of the first generation of the broadband internet and the available broadband internet speed in later years when the technology was improved. Whereas Falck et al. (2014) relate the distance between the geographical middle point of a municipality to the threshold, we refine their strategy by calculating and using the distance of each establishment to the MDF.

We exploit the fact that distance to the MDF mattered to get access to higher broadband internet speed in Germany. We compare establishments with differing distance to their respective MDFs and follow them over time. Our strategy allows ruling out time-constant characteristics of the establishments or firms by controlling for the respective fixed-effects. We further control for county-year characteristics to take potential changes in the labor market into account.

We assemble a unique dataset for our analysis. We use German social security records that provide information on establishments and employees. We combine the administrative data with firm-level information by Bureau van Dijk via a record linkage procedure. Our data thus allow us to see which establishments belong to the same firm. We geocode the establishments and add information on the telephone network to assess which broadband internet speed is available to the establishment.

In a first step, we study changes in employment at the establishment and firm level. In particular, we analyze how the availability of faster broadband internet to one establishment impacts its own employment as well as the employment at other establishments of the same firm. We study the impact of the available broadband internet speed on the subordinate establishment, HQ, and firm.

First, we study the impact of the internet speed available to the subordinate establishments and its HQ on the employment in the subordinate establishment. We find that establishments with access to faster broadband internet grow more slowly. However, establishments in firms in which the HQ has access to faster internet grow faster. Second, we study the impact of the broadband internet speed of the HQ as well as the average internet speed available to the subordinate establishments of the firm on the employment in the HQ. We do not find significant effects. Third, we use the same regression equation as for the HQ to study the impact on the whole firm. Again, we find no sizeable effects. Overall, our results suggest an interdependence of the subordinate establishments and the HQ within a firm as subordinate establishments respond to their HQ's access to faster broadband internet. However, this interdependence seems to mainly affect employment at the subordinate establishment.

In a second step, we study the impact of the availability of broadband internet speed on the skill composition at the establishment and firm level. Subordinate establishments with access to faster internet increase the share of low-skilled labor but decrease the share of high-skilled labor. However, the skill composition of subordinate establishments also depends on the access to faster broadband internet of the HQs of their firms. Subordinate establishments of firms in which the HQ gets access to faster broadband internet decrease the share of low-skilled labor.

Similar to subordinate establishments, HQs also increase the share of low-skilled labor and decrease the share of high-skilled labor when they get access to faster broadband internet. However, the average available broadband internet speed of the respective subordinate establishments does not affect the skill composition of the HQ. Further, the share of low-skilled employment of the whole firm increases in firms whose HQ gets access to faster broadband internet. The share of high-skilled employment in these firms decreases. The average available speed of the broadband internet of the subordinate establishments does not have affect the skill composition of the firm.

Again, our results suggest an interdependence of the establishments affecting the skill composition of the subordinate establishment. Subordinate establishments respond to their HQ's access to faster broadband internet. However, we do not find any evidence that the interdependence of subordinate establishments and HQs affects the skill composition at the HQ. Further, our results for the response of subordinate establishments to their own access to faster broadband internet, in contrast to the response to their HQ's access, is at odds with previous findings in the literature (Akerman et al., 2015). To interpret these results, one should, however, keep in mind that the establishments of multi-establishment firms may not be equal when faster broadband internet becomes available. They may rather fulfill different functions within the firm. To account for this possibility, we split our sample of subordinate establishments along the initial skill composition, i.e., by positive or zero employment of high-skilled labor in 1999.

We find that the effects on employment are driven by establishments that employed high-skilled labor before the introduction of the broadband internet. Subordinate establishments with high-skilled employment grow more slowly when they get access to faster broadband internet. Further, they grow faster when the HQ gets access to faster broadband internet. We find no significant effects on subordinate establishments that did not employ high-skilled labor before the introduction of broadband internet. This finding suggests that establishments that provide the complementary input respond more strongly to the technology than those establishments that did not employ high-skilled labor. It is thus consistent with the idea of skill-complementarity of the broadband internet.

Further, subordinate establishments increase the shares of more skilled labor when the HQ gets access to faster broadband internet. Subordinate establishments that employ high-skilled labor and whose HQ gets access to faster broadband internet reduce the share of medium-skilled and increase the share of high-skilled labor. Subordinate establishments that do not employ high-skilled labor decrease the share of low-skilled labor and increase the share of medium-skilled labor with the HQ's access to faster broadband internet.

Overall, we find that subordinate establishments do not only respond to their own access to broadband internet but also to their HQ's access. This finding suggests an interdependence of the establishments. Further, we find that the effects differ depending on the initial skill composition of the subordinate establishments. Subordinate establishments that already employ high-skilled labor respond more strongly. Besides, they shift their skill composition towards more skilled labor when their HQ gets access to faster internet. We thus provide supportive evidence for skill-complementarity of faster broadband internet.

To the best of our knowledge, we are the first to study the impact of the access to higher broadband internet speed on the employment and skill-composition of multi-establishment firms. We contribute to three strands of literature. First, we build on and extend the literature on the impact of ICT, in particular broadband internet, on firm behavior.[2] The existing literature finds that technology may change optimal firm boundaries (Baker and Hubbard, 2003; Baker and Hubbard, 2004; Acemoglu et al., 2010). Fort (2017) provides empirical evidence suggesting that communication technology lowers

[2]For a detailed survey of the literature on the impact of ICT on productivity we refer to Draca et al. (2006) and Cardona et al. (2013). The literature on broadband internet is surveyed by Bertschek et al. (2015).

coordination costs as firms adopting the technologies are more likely to fragment their production.

Akerman et al. (2015) find positive output elasticities as well as skill-biased employment and wage effects of the broadband internet. They exploit the quasi-exogenous time variation in broadband infrastructure expansion in Norwegian municipalities due to limited funding of a government initiative. This chapter adds to their findings in a number of ways. First, we study the impact of the availability of different broadband internet speed on multi-establishment firms. Second, we look at within-establishment responses, which allow us to control for any time-constant establishment characteristics. Further, we contribute to this literature by showing that our results on employment are driven by establishments that already employed high-skilled labor. Moreover, they increase the share of skilled labor when getting access to faster broadband internet, suggesting specialization of the establishments.

Second, we extend the literature on multi-establishment firms. Gumpert et al. (2018) study how geographic frictions affect firms' managerial organization. They show that geographic frictions between the HQ and one subordinate establishment affect the organization of all establishments of a multi-establishment firm. Similarly, Charnoz et al. (2015) show empirically that high-speed train routes decrease the share of managers at subordinate establishments that are affected by lower travel times to the HQ. We contribute to this literature by showing how ICT affects multi-establishment firm employment and skill composition. Further, we study how the access to faster internet of one establishment affects the other establishments of the firm.

Third, we extend the instrumental variable strategy by Falck et al. (2014) who study voting behavior. They exploit the distance between the geographical middle point of a municipality to its MDF as an instrument for the share of households in a municipality with broadband access. We build on this idea by calculating the distance between each individual establishment and the dedicated MDF. Hence, we approximate the access to faster broadband internet at the establishment-level instead of the municipality-level. Our strategy allows measuring the technological restriction more precisely and calculating the intention-to-treat effect on similar establishments within the same municipality.

The chapter is structured as follows. The next section describes our empirical strategy. Section 2.3 describes the data we use. In section 2.4, we present our findings. Section 2.5 concludes.

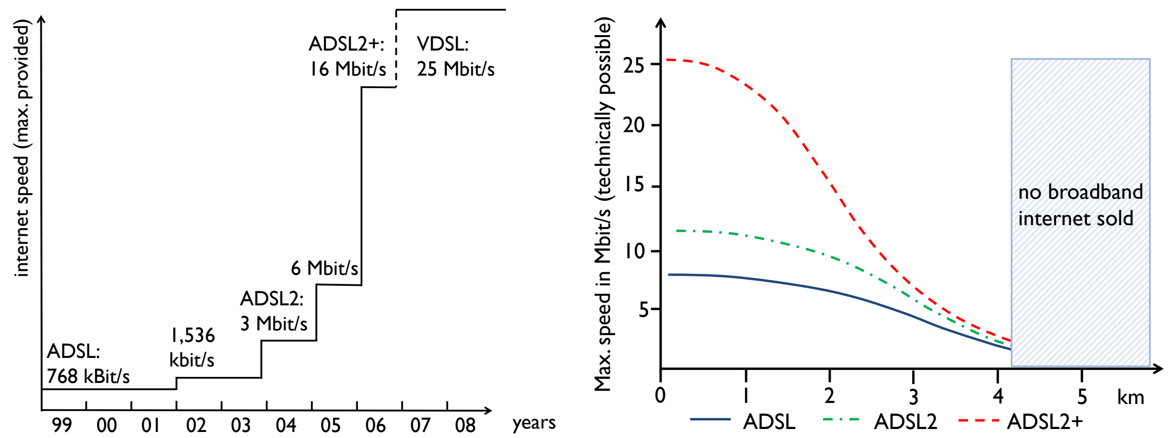
2.2 Empirical strategy

We are interested in the impact of the availability of faster broadband internet at the establishment level on the employment and skill composition of multi-establishment firms. In our regression equation, we distinguish between the broadband internet speed available to the HQ and the subordinate establishments of the firm. If we observed broadband internet speed adoption, a naive regression would consist of the following equation:

$$
\begin{aligned}
y_{jt} =& \beta_1 \times \textsf{HQ adopts broadband internet speed}_{i,t} \\
&+\beta_2 \times \textsf{subordinate establishment adopts broadband internet speed}_{j,t} \\
&+\alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}
\end{aligned}
$$

In this regression, we would regress e.g. employment y_{jt} of subordinate establishment j in year t on both a dummy equal to one if the HQ of firm i adopts a certain broadband internet speed and a dummy equal to one if the subordinate establishment itself adopts a certain broadband internet speed. This specification would suffer from omitted variable bias. Early-adopting firms are likely to be different in terms of other potentially time-varying characteristics as firms self-select into adopting more advanced internet technologies. For example, early-adopting firms may also be more innovative and choose to adopt the internet in order to bring their new products to a bigger market. Thus, we would overestimate the impact of internet speed on any outcome which is positively correlated with innovative activities.

Furthermore, a firm could decide to adopt different speeds in each establishment depending on the intended use of the broadband internet. As a result, different effects of higher internet speeds by establishment type may be driven by endogenous choices of the firm to only adopt higher speeds where they are efficient. If we could conduct an ideal experiment, we would randomize internet speed to firms and even to different establishments within firms. In practice, randomization is not feasible.



(a) Timeline (b) Decay of the broadband internet speed

Figure 2.1: Illustration of identification strategy

The left figure shows the timeline of the introduction and upgrading of the broadband internet technology following Schnabel (2015). In 2000, customers received ADSL with a maximum download speed of 768 kBit/s. In later years, the speed was upgraded up to 6 Mbit/s. In 2006, however, ADSL2+ provided a much larger speed upgrade with up to 16 Mbit/s. Further, VDSL was introduced but required large infrastructure investments which took several years.
The right figure shows the decay of the maximum broadband internet speed that was technically possible with increasing distance to the MDF. The solid line shows the decay for the first generation (ADSL). The line above shows the decay for less used technology called ADSL2. The dashed line shows the decay for the second major generation of the broadband internet (ADSL2+). Source: Schnabel (2015) and own illustrations.

To overcome this problem, we exploit the fact that technological peculiarities of the traditional public switched telephone network (PSTN) led to differential availability of broadband internet speed levels (Falck et al., 2014). When the broadband internet was introduced in the 2000s, it was transferred through the copper wires of the existing telephone network. The length of the copper wire, and hence the distance, from an establishment to the respective MDF is decisive for the internet speed an establishment could receive.

In our analysis, we exploit the variation in internet speed over time. As shown in Figure A.3, broadband internet speed increased in the 2000s (see Schnabel, 2015, for more details). The first generation of broadband internet, asymmetric digital subsriber line (ADSL), was introduced at a fair in 1999. In 2000, 600,000 customers subscribed to the new technology as shown in figure B.1 in the appendix. Download speed and broadband internet adoption increased in later years. In 2002, it was upgraded to 1,536 kBit/s. In 2004, ADSL2 was introduced reaching up to 3 Mbit/s and upgraded to 6 Mbit/s in 2005. From 2006 on, ADSL2+ was introduced providing up to 16 Mbit/s by most telecommunications providers.

Figure 2.1b shows the decay in download internet speed with increasing distance to the MDF. Establishments located up to 4.2 km from the MDF could receive the first generation of broadband internet. The state monopolist at the time of the introduction of the broadband internet did not sell broadband internet to establishments with longer copper wires as they could not provide their minimum standards of speed. The provided internet speed of the first generation of broadband internet and the later speed upgrades decay with the distance to the MDF. The line above shows the decay of broadband internet speed of ADSL2 which starts at a higher speed but decays even more steeply. ADSL2+ is considered the second major generation of the broadband internet. The decay in internet speed is most pronounced for this generation. Establishments located very close to the MDF could theoretically receive up to 25 Mbit/s but most telecommunications providers only sold up to 16 Mbit/s. Establishments located 2 km from the MDF would receive up to 16 Mbit/s, whereas establishments located 3 km from MDF would hardly notice a speed upgrade compared to previous generations.

The PSTN was installed in the 1960s by the state monopolist at that time. The infrastructure consists of a large backbone system which is connected to around 8,000 MDFs in Germany. The length of the copper wires between an establishment and its MDF does not affect the quality of telephone communications. It is therefore unlikely that establishments located strategically close to the MDF before the introduction of the broadband internet.

To establish our identification strategy, we combine the exogenous geographical variation in access to faster broadband internet with the timing of the speed upgrades. The distance to the MDF determines the treatment intensity of high-speed internet. Our main regression equation takes both dimensions into account and estimates the intention-to-treat effect:

$$\begin{aligned} y_{jt} = &\beta_{1,t} \times \log(\textsf{HQ distance to MDF})_i \times Period_t \\ &+\beta_{2,t} \times \log(\textsf{subordinate establishment distance to MDF})_j \times Period_t \\ &+\alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}, \end{aligned} \tag{2.1}$$

We interact the distance of the subordinate establishment j or the headquarters of firm i respectively with dummies for three different time periods. We use 1999 as our reference year in which no firm had access to faster broadband internet. Our first time period is defined for the years from 2000 to 2003. Establishments could receive up to 1,536 kBit/s in download speed. Our second time period ranges from 2004 to 2005 in which the maximum provided speed increased to up to 6 Mbit/s. Our third time period from 2006 to 2010 covers the years in which internet speed was increased due to the second major generation of ADSL2+. We control for subordinate establishment-fixed effects α_j to rule out time-constant subordinate establishment characteristics. Further, we compare subordinate establishments within the same county in the same year with differing broadband internet speed availability by controlling for county-year fixed effects α_{ct}. $\varepsilon_{i(j)t}$ is the error term clustered at the county-year-level.

We further run regressions at the HQ-level and the firm-level. In both cases, we use the interaction terms of the distance of the HQ to the MDF with time period dummies as described above. To take the average available internet speed of the subordinate establishments into account, we calculate the average distance of the subordinate establishments to their MDFs, weighted by employment.[3] We interact the weighted average distance with time period dummies.

Distance to the MDF is negatively correlated with internet speed up to a certain threshold. We restrict our sample to establishments located up to six km from the MDF. Above this threshold, distance to the MDF does not affect the internet speed as the broadband internet cannot be transmitted. Additionally, this restriction allows us to exclude firms that may be located far away from economic centers for other reasons and may hence evolve differently over time.

[3]Chapter 1 uses dummies comparing single-establishment firms located below and above a threshold for certain internet speeds. To reduce measurement error, we leave out "donut" holes in chapter 1. We choose distance to the MDF as our explanatory variable as we need to aggregate the internet speed available to the subordinate establishments. Using the approach by chapter 1 would not allow aggregating the information especially considering the "donut" holes.

2.3 Data and descriptive statistics

We assemble a unique longitudinal firm-establishment-employee-level dataset. We combine three data sources: the German social security data, the ORBIS database by Bureau van Dijk, and geocoded information on the telephone network. For our analysis, we use an unbalanced panel from 1999 to 2010. We restrict our sample to multi-establishment firms.

Firm-establishment-employee data. The German social security records contain information on establishments and all of their employees subject to social security contributions.[4] We have information on the county, age, and three-digit sector of each establishment. We keep establishments that were founded before the introduction of the broadband internet. For each employee, we observe the age, gender, education, occupation, employment history, and wages. We impute missings in the education variable following Fitzenberger et al. (2005)[5]. We restrict our sample to full-time employees. We combine the social security data with balance sheet information in the ORBIS database provided by Bureau van Dijk. We employ a record linkage procedure to allocate establishments to firms described in Appendix C.1.1. Each firm, establishment, and employee is allocated a unique identifier that allows following each unit of observation over time.

Information on the broadband internet speed availability. To approximate individual broadband availability, we calculate the distance between each establishment to its dedicated MDF. We geocode each establishment based on the address included in the social security records. We allocate each establishment to its local network using the information on geocoordinates of the telephone network from Bundesnetzagentur (2017). We use the closest main distribution frame in this local network if there are several MDFs in the network. If there is no MDF in the local network, we use the closest one outside the borders. We calculate the distance from each establishment to the MDF via roads. This distance is more accurate than taking the airline distance as the cables

[4] We use the same dataset as in chapter 3. To make each chapter self-contained, we describe the data for each chapter separately.

[5] After the imputation, we still have missings in 2.5% of the observations. We impute these remaining observations by transferring the dominant educational background in each occupation.

are installed belowground. Opening the ground next to roads was the cheapest way to install the cables. We use the geocoded information on the road network provided by OpenStreetMap (2017). The distance is calculated based on the cross-sectional information provided in 2015 and does not vary over time. We drop around 15% of the establishments in our sample for which we cannot calculate the distance to the MDF due to missing or incomplete reporting of the address.

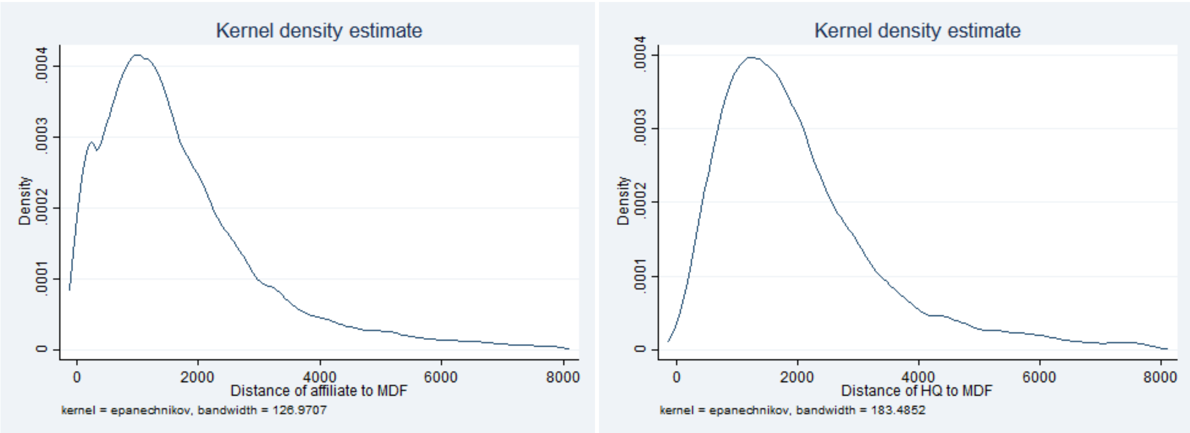
Outcome variables. We are interested in the employment effects of the broadband internet availability to study growth and labor allocation within the firm. We use full-time employment at the establishment and firm level as our outcome variables. Further, we study the skill composition to understand which skill groups are most affected. We define three skill groups based on the information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training or a university-entrance qualification. High-skilled employees have at least a college degree.

Descriptive statistics. Table 2.1 provides descriptive statistics for our data set. Our sample comprises around 5,300 multi-establishment firms. We observe each firm for 10.6 years on average. Firms have around six establishments and 300 employees on average. The distribution of both employment and number of establishments is highly skewed. The median firm has two establishments and 56 employees. Headquarters (HQs) have 120 employees and represent 58% of firm employment on average. Our sample includes around 29,000 subordinate establishments. As firms add and drop establishments, the average number of years per subordinate establishment is lower than per firm. We observe each subordinate establishment for 9.2 years on average.

Low-skilled employees represent the smallest shares of employment at the firm, HQ, and subordinate establishment level. On average, around 8% of employees are low-skilled. The median subordinate establishment does not even employ any low-skilled labor. The median HQ and the median firm employ 2.3% of low-skilled labor.

Most firms, HQs, and subordinate establishments employ medium-skilled labor and it is usually the largest group at the establishment and firm level. The average share of medium-skilled employment ranges from 78.6% at the HQ level to 84.2% at the subor-

Figure 2.2: Density of distance to MDF by establishment type



The figure shows the kernel density of the distance to the MDF in meters. The left graph shows the kernel density for subordinate establishments. The right graph shows the kernel density for headquarters.

dinate establishment level. The median shares are even higher, ranging up to 100% at the subordinate establishment level.

High-skilled labor makes a larger share of employment than low-skilled labor on average. Especially at the HQ level, the share of high-skilled labor is higher on average and at the median. At the subordinate establishment level, however, the average share of high-skilled is only around two percentage points higher than the share of low-skilled labor. Only around one-third of subordinate establishments employ any high-skilled labor.

The distribution of the distance from the HQs and subordinate establishments to the MDF is skewed to the left. On average, subordinate establishments are located around 1,700 m from the MDF. The median subordinate establishment is located around 1,300 m from the MDF. The average distance to the MDF of HQs is around 2,000 m, whereas the median HQ is located around 1,700 m from the MDF. Figure 2.2 shows the kernel densities of distance to the MDF by establishment type. The distributions of the two different establishment types are very similar. A larger share of subordinate establishments is located closer to the MDF.

Table 2.1: Descriptive statistics

Units of observation	N	Unique IDs					
Firms	56,885	5,370					
Headquarters	55,047	5,379					
Sub-establishments	223,261	29,083					
Employees	16,992,391	2,959,314					
Descriptive statistics	**N**	**Mean**	**SD**	**p25**	**p50**	**p75**	**p95**
Firms							
Employment	55,151	305	2,852	22	56	158	902
Number of establishments	55,151	5.6	40.1	2	2	4	55
Shares (in %)							
Low-skilled	55,151	7.5	11.7	0	2.3	10.4	32.1
Medium-skilled	55,151	79.1	19.7	71.0	84.2	93.3	100
High-skilled	55,151	11.1	17.5	0	3.9	14.0	52.6
Headquarters							
Employment	55,047	121.2	473.7	10	30	91	463
Shares (in %)							
In firm empl.	55,047	58.5	26.4	38.6	61.5	80.6	96
Low-skilled	55,047	8.0	13.0	0	2.3	10.7	35.2
Medium-skilled	55,047	78.6	21.2	68.2	83.9	95.5	100
High-skilled	55,047	13.4	19.9	0	5	18.2	58.8
Distance to MDF (in m)	5,213	1,911	1,168	1,060	1,676	2,497	4,347
Subordinate establishments							
Employment	223,261	38.6	368	2	6	17	116
Shares (in %)							
In firm empl.	223,261	12.9	20.1	.4	3.5	16.7	60
Low-skilled	223,261	6.7	15.6	0	0	5.3	37.5
Medium-skilled	223,261	84.2	24.7	77.1	100	100	100
High-skilled	223,261	9.1	21.2	0	0	5.6	60.6
Distance to MDF (in m)	29,083	1,612	1,195	756	1,352	2,183	4,109

This table shows the summary statistics by firm, HQ, and subordinate establishment. The samples for HQs and subordinate establishments are restricted to establishments within six km to the MDF. The sample of firms is not restricted. The information on skills reported is social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training or a university-entrance qualification. High-skilled employees have at least a college degree. Shares are reported in percent.

2.4 Results

This section presents results from estimating equation 2.1 for subordinate establishments and the adjusted equations for HQs and firms. In the appendix, we show the results from several robustness checks as well as yearly effects instead of aggregated time periods.

Employment. Table 2.2 presents the results on employment. The first column shows the results from estimating equation 2.1 for subordinate establishments. It shows that subordinate establishments that are closer to the MDF, i.e. with faster internet available, grow more slowly. During the time period from 2006 to 2010, a one percent increase in distance to the MDF is associated with 0.02 percent less employment (significant on the 1%-level). Hence, a subordinate establishment at the first quartile of the distance distribution would be 4.3 percent smaller in the time period from 2006 to 2010 than a subordinate establishment at the third quartile. Further, subordinate establishments of firms with access to faster internet speed in the HQ grow faster. A subordinate establishment with an HQ located at the first quartile of the distance distribution is 5.6 percent larger during the time period from 2006 to 2010 than if the HQ was located at the third quartile. The effect kicks in during the time period from 2004 to 2005 but the magnitudes of the coefficients increase in the later period.

The second column shows the employment results for the HQ. The HQ's employment does not respond significantly to the access to higher internet speed. The magnitude of the coefficient for the time period 2006 to 2010 is less than one-fifth of the coefficient for the subordinate establishment. We find no effect of the average available internet speed of the subordinate establishments. The third column shows the results for the whole firm. We find no effect of the HQ's access to faster broadband internet or the average internet speed available to the subordinate establishments.

The employment results suggest that access to higher internet speed matters for multi-establishment firm employment. Subordinate establishments grow more slowly with increasing own internet speed available and faster with the HQ's internet speed available. This is suggestive evidence for an interdependence of the subordinate establishment and the HQ. However, this interdependence does seem to affect the employment at the HQ.

Table 2.2: Regression results on employment

	Sub-estab. (1)	HQ (2)	Firm (3)
Sub-estab. distance × $D_{2000-03}$	0.002		
	(0.005)		
Sub-estab. distance × $D_{2004-05}$	0.011^{+}		
	(0.007)		
Sub-estab. distance × $D_{2006-10}$	0.022^{**}		
	(0.008)		
HQ distance × $D_{2000-03}$	−0.007	0.005^{+}	0.004^{+}
	(0.005)	(0.003)	(0.002)
HQ distance × $D_{2004-05}$	-0.025^{***}	0.001	0.000
	(0.007)	(0.004)	(0.003)
HQ distance × $D_{2006-10}$	-0.039^{***}	0.007	0.006^{+}
	(0.007)	(0.005)	(0.003)
Avg. sub-estab. distance × $D_{2000-03}$		−0.001	0.000
		(0.004)	(0.003)
Avg. sub-estab. distance × $D_{2004-05}$		0.001	0.002
		(0.004)	(0.003)
Avg. sub-estab. Distance × $D_{2006-10}$		−0.003	0.000
		(0.005)	(0.003)
R-squared	0.902	0.959	0.972
Obs.	217,387	55,047	55,151

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. For the estimations at the HQ- and firm-level, the average distance of the subordinate establishments of the firm to their respective MDFs is used (weighted by employment). The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Skill composition. Table 2.3 reports the results on the skill composition. We report three different outcome variables representing the skillshares. Columns one to three report the results at the subordinate establishment level. Subordinate establishments with faster internet available increase the share of low-skilled employment and decrease the share of high-skilled employment. A subordinate establishment at the first quartile of the distribution of the distance to the MDF employs a 0.75 percentage points higher share of low-skilled labor and 0.51 percentage points lower share of high-skilled labor during the time period from 2006 to 2010 than a subordinate establishment at the third quartile. Compared to the average shares of low- and high-skilled employment, this result is sizeable. It represents 11% (6%) of the average shares of low-skilled (high-skilled) employment. The effect on the share of medium-skilled employment is hardly significant but suggests a decrease. The skillshares of subordinate establishments also respond to the available internet speed in the HQ. Subordinate establishments in firms with faster internet speed available at the HQ decrease the share of low-skilled employment. We find no effect on medium-skilled and high-skilled employment.

Columns four to six report the results for HQs. HQs with access to faster internet increase the share of low-skilled employment and decrease the share of high-skilled employment. We find no significant effect on the share of medium-skilled employment. An HQ at the first quartile of the distance distribution employs a 0.25 percentage points higher share of low-skilled labor and 0.28 percentage points lower share of high-skilled labor than an HQ at the third quartile from 2006 to 2010. The effect already seems to start in smaller magnitudes during the time period from 2000 to 2003 and then fully kicks in from 2004 to 2005. Our results suggest that the skillshares at the HQ of a firm do not change with the average available internet speed at the subordinate establishments.

Columns seven to nine report the results for the whole firm. Firms with an HQ that gets access to faster internet increase the share of low-skilled employment and decrease the share of medium- and high-skilled employment. A firm whose HQ is located at the first quartile of the distance distribution employs a 0.27 percentage points higher share of low-skilled labor from 2006 to 2010 than a firm whose HQ is located at the third quartile. The skill composition of the firm hardly changes with the average internet speed available to its subordinate establishments. The results suggest that firms decrease the share of low-skilled employment and increase the share of medium-skilled employment but the coefficients are not significant.

The results on the skill composition suggest that firms change the skill composition at their HQ and subordinate establishments when getting access to faster internet. The increases of the shares of low-skilled employment at the subordinate establishments and HQs with their own available internet speed, however, contradict previous work finding skill complementarity of the broadband internet. However, responses of subordinate establishments and HQs to their respective counterpart's internet speed available counteract the response to the own internet speed. This finding again suggests an interdependence of the subordinate establishments and HQs of a multi-establishment firm. This interdependence affects the skill composition at the subordinate establishment.

Differential effects by initial skill composition. Depending on the internal allocation of the skilled workforce, multi-establishment firms may differently adjust their workforce as a response to faster broadband internet availability. In fact, only about 25% of subordinate establishments employ all three skill groups but almost all subordinate establishments employ medium-skilled labor. Hence, it may be important to take the initial skill composition of establishments into account. We split our sample of subordinate establishments by positive or zero employment of high-skilled labor in 1999. We analyze and report the changes in the skill groups that were initially most relevant for the respective subordinate establishments.[6]

Table 2.4 reports the results for the two separate samples. The left panel shows the results for subordinate establishments that employed high-skilled labor in 1999. Column one shows the results on employment. The results are similar to the results reported in table 2.2 but larger in magnitude. The effect starts during the time period from 2000 to 2003 and becomes stronger over time. Subordinate establishments with access to faster internet grow more slowly. A subordinate establishment located at the first quartile of the distance distribution is about 12% percent smaller during the time period from 2006 to 2010 than a subordinate establishment at the third quartile. However,

[6]We report the remaining results in table B.1 in Appendix B.2.1. We do not interpret the results as they are not meaningful for our analysis. For example, only 15% of subordinate establishments that did not employ high-skilled labor in 1999 ever employ high-skilled labor till 2010.

Table 2.3: Regression results on the skill composition

	Sub. establishment			HQ			Firm		
	Low (1)	Medium (2)	High (3)	Low (4)	Medium (5)	High (6)	Low (7)	Medium (8)	High (9)
Sub-e. distance × $D_{2000-03}$	−0.126	0.113	0.012						
	(0.081)	(0.095)	(0.063)						
Sub-e. distance × $D_{2004-05}$	−0.306**	0.191+	0.115						
	(0.111)	(0.144)	(0.098)						
Sub-e. distance × $D_{2006-10}$	−0.380***	0.120	0.260*						
	(0.114)	(0.151)	(0.106)						
HQ distance × $D_{2000-03}$	0.098	−0.121	0.023	−0.090**	0.064	0.026	−0.056*	0.057	0.048+
	(0.079)	(0.092)	(0.061)	(0.032)	(0.047)	(0.041)	(0.028)	(0.042)	(0.029)
HQ distance × $D_{2004-05}$	0.232*	−0.180+	−0.052	−0.168***	0.014	0.153**	−0.166***	0.123*	0.098**
	(0.110)	(0.139)	(0.093)	(0.042)	(0.058)	(0.051)	(0.039)	(0.053)	(0.035)
HQ distance × $D_{2006-10}$	0.271*	−0.118	−0.154	−0.176***	−0.021	0.197***	−0.187***	0.111	0.143**
	(0.114)	(0.146)	(0.101)	(0.051)	(0.070)	(0.058)	(0.044)	(0.069)	(0.045)
Avg. sub-e. distance × $D_{2000-03}$				0.052	−0.056	0.004	0.016	−0.014	−0.032
				(0.033)	(0.050)	(0.045)	(0.031)	(0.046)	(0.032)
Avg. sub-e. distance × $D_{2004-05}$				0.062	−0.006	−0.056	0.075+	−0.077	−0.029
				(0.044)	(0.062)	(0.056)	(0.044)	(0.057)	(0.039)
Avg. sub-e. distance × $D_{2006-10}$				0.013	0.002	−0.014	0.042	−0.071	0.010
				(0.049)	(0.071)	(0.062)	(0.047)	(0.070)	(0.048)
R-squared	0.759	0.828	0.884	0.866	0.888	0.919	0.902	0.903	0.949
Obs.	223,261	223,261	223,261	55,047	55,047	55,047	55,151	55,151	55,151

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. For the estimations at the HQ- and firm-level, the average distance of the subordinate establishments of the firm to their respective MDFs is used (weighted by employment). The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training or a university-entrance qualification. High-skilled employees have at least a college degree.

subordinate establishments in firms in which the HQ gets access to faster internet grow faster. A subordinate establishment whose HQ is located at the first quartile of the distance distribution is 14% larger than a subordinate establishment whose HQ is located at the third quartile.

Column two and three report the results on the shares of medium- and high-skilled employment. The internet speed available to the subordinate establishment itself does not have a significant effect on the skillshares. However, the skillshares in the subordinate establishment respond to the available internet speed at the HQ. Subordinate establishments whose HQ gets access to faster internet decrease the share of medium-skilled employment and increase the share of high-skilled employment. A subordinate establishment whose HQ is located at the first quartile of the distance distribution employs 0.9 percentage points less medium-skilled and more high-skilled labor from 2004 on. The average subordinate establishment in this subgroup employs about 23% high-skilled employment. Hence, this change represents about 4% of the mean. The effect kicks in during the time period from 2004 to 2005.

The right panel shows the results for subordinate establishments that did not employ any high-skilled labor in 1999. We find no significant effect of the internet speed available to the subordinate establishment or the HQ on employment. Regarding the skill composition, we find that subordinate establishments that get access to higher internet speed increase the share of low-skilled employment. A subordinate establishment at the first quartile of the distance distribution employs about one percentage point more low-skilled labor from 2004 on than a subordinate establishment at the third quartile. We find no significant effect on the share of medium-skilled labor. Subordinate establishments in this subsample also respond to the available internet speed at the HQ. Faster internet available at the HQ reduces the share of low-skilled and increases the share of medium-skilled employment. A subordinate establishment of a firm whose HQ is located at the first quartile of the distance distribution employs about 0.5 percentage points less low-skilled and more medium-skilled labor from 2004 on. The effect kicks in during the time period from 2004 to 2005.

Our results suggest differential responses of subordinate establishments to access to faster internet by their initial skill composition. Our findings on employment in table 2.2 are driven by subordinate establishments that employed high-skill labor in 1999. This finding is in line with previous work pointing towards skill complementarity of the broadband internet as subordinate establishments that provide complementary labor respond more strongly to internet speed access.

The changes in the skill composition also differ by the two samples. Subordinate establishments that did not employ high-skilled labor in 1999 increase the share of low-skilled labor when getting access to faster internet. However, subordinate establishments also respond significantly to the available internet speed at the HQ. Depending on the initial skill composition, they decrease the share of less-skilled employment and increase the share of the more skilled employment when the HQ gets access to faster internet. This finding is in line with both skill complementarity and the interdependence of the establishments in multi-establishment firms.

Considering the timing of the effects, our results suggest that the effects are not driven by the latest maximum speed upgrades. The effect tends to kick in during the time period from 2004 to 2005 already. Hence, it seems to be driven by earlier speed upgrades before the large upgrade to the second major generation.

Robustness checks. To rule out any time-varying differences of subordinate establishments that correlate with the distance to the MDF, we run several robustness checks. In sub-section B.2.3 in the appendix we report the robustness checks on employment (table B.5), the share of low-skilled (table B.6), medium-skilled (table B.7), and high-skilled (table B.8) employment at the subordinate establishment level.

First, one may be worried that firms strategically locate close to the MDF, or the MDF may be installed close to specific firms that grow and adjust their skill composition differently. As explained above, the network in Western Germany was installed in the 1960s. Hence, MDFs could be located close to firms that existed at that time. As a robustness check, we run separate regressions on firms founded before the 1990s when first internet technologies were introduced. Second, we run separate regressions excluding Eastern Germany where the MDFs were installed only in the 1990s.

Third, we check for potential measurement error because of lagged technology upgrading. If the MDF did not provide broadband internet, e.g., because it was not up-

Table 2.4: Regression results by initial skill composition

	With high-skilled empl. in 1999			Without high-skilled empl. in 1999		
		Skillshares			Skillshares	
	Employment (1)	Medium (2)	High (3)	Employment (4)	Low (5)	Medium (6)
Sub-estab. distance × $D_{2000-03}$	0.030***	0.005	0.057	0.010*	−0.112	−0.061
	(0.011)	(0.214)	(0.215)	(0.005)	(0.103)	(0.113)
Sub-estab. distance × $D_{2004-05}$	0.041**	−0.302	0.412	0.006	−0.460***	0.173
	(0.018)	(0.321)	(0.318)	(0.009)	(0.170)	(0.187)
Sub-estab. distance × $D_{2006-10}$	0.060***	−0.307	0.430	0.006	−0.486***	0.081
	(0.020)	(0.321)	(0.306)	(0.010)	(0.175)	(0.201)
HQ distance × $D_{2000-03}$	−0.044***	0.245	−0.274	−0.008	0.095	−0.068
	(0.011)	(0.215)	(0.215)	(0.005)	(0.099)	(0.109)
HQ distance × $D_{2004-05}$	−0.072***	0.634**	−0.666**	−0.012	0.391**	−0.328*
	(0.018)	(0.319)	(0.316)	(0.008)	(0.168)	(0.182)
HQ distance × $D_{2006-10}$	−0.096***	0.633**	−0.648**	−0.015	0.356**	−0.244
	(0.020)	(0.316)	(0.301)	(0.009)	(0.175)	(0.197)
R-squared	0.906	0.876	0.895	0.877	0.757	0.726
Obs.	51,239	51,239	51,239	103,518	103,518	103,518

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training or a university-entrance qualification. High-skilled employees have at least a college degree. Samples are split by zero or positive employment of high-skilled labor in 1999.

graded to provide the new technology yet, establishments around the MDF would falsely be assumed to receive broadband internet. Hence, we use information on the share of households with broadband internet from 2005 to 2008 provided by the Federal Ministry for Economic Affairs and Energy (Federal Ministry of Economics and Technology, 2009). As a robustness check, we exclude municipalities in which no household had broadband internet access in 2005. Fourth, our estimates may be biased by the introduction of Very High Speed Digital Subscriber Line (VDSL). Therefore, we exclude all counties in which VDSL was introduced until 2008.

The results show that our findings are robust to all four changes to the specification as reported in the four columns of each table for the respective outcome variables. The sizes of most of the coefficients hardly change, even though the magnitude tends to be higher when excluding counties in which VDSL was introduced until 2008.

To point out the role of available internet speed in comparison to the mere access to the broadband internet, we run regressions excluding subordinate establishments and HQs located further than 4.2 km from the MDF. We report the results for the pooled sample in table B.9. We find that the results are very similar. If anything, the effects on employment are even stronger than for the main sample. Moreover, we split the sample by positive and zero high-skilled employment in 1999 as in table 2.4. We report the results in table B.10. We find very similar results. Further, the available internet speed at the HQ has a significantly positive effect on employment growth for subordinate establishments without high-skilled labor in 1999 in this subsample.

Previous literature finds differential effects for non-manufacturing and manufacturing firms. To contribute to this discussion, we split our sample by the sector reported for the respective HQ. We report the results for non-manufacturing firms in table B.11. We find similar effects as for the pooled sample. In addition, we split the sample by the initial skill composition as in table 2.4. The results are reported in table B.14. We find very similar results as in the main specification. Besides, the available internet speed at the HQ has a significantly positive effect on employment growth for subordinate establishments without high-skilled labor in 1999 in this subsample.

For manufacturing firms, we report the results in table B.13 for the pooled sample in the appendix. We find no effect of the broadband internet speed availability on employment and the skill composition in the pooled sample. Table B.12 reports the results for the sample split by initial skill composition. The results reveal counteracting effects on employment. The available internet speed at the HQ has a positive effect on employment growth for subordinate establishments with high-skilled labor in 1999 but a negative effect for those without high-skilled labor in 1999. Overall, our main findings seem to be driven by non-manufacturing firms.

2.5 Conclusion

This chapter provides new insights on the impact of the broadband internet availability on the employment and skill composition of multi-establishment firms. We show that access to faster broadband internet has a significant impact on the employment at subordinate establishments and the skill composition of subordinate establishments and HQs. Moreover, subordinate establishments do not only respond to their own broadband internet availability but also to the availability at their HQ's location. Our findings suggest both interdependence of subordinate establishments and HQs as well as skill complementarity of the broadband internet. Multi-establishment firms are key players in today's value chains and employment allocation. Hence, the impact of new technologies on these firms remain an interesting area for future research.

Chapter 3

Firm Organization with Multiple Establishments

3.1 Introduction

Large firms often organize their employees in multiple establishments at different locations. Geographic frictions between the subordinate establishments and the headquarters, such as higher distance or longer travel times between their locations, adversely affect the performance of the subordinate establishments (e.g., Giroud, 2013; Kalnins and Lafontaine, 2013). Anecdotal evidence suggests that adjusting the managerial organization may help firms mitigate the negative impact of geographic frictions. For example, employing middle managers at regional offices instead of at the headquarters was a key ingredient for the success of Singer Sewing Machine in the US (Chandler, 2002, p. 403-5). Philips employed dedicated country managers and regional executives as part of a larger strategy to revitalize their operations after 1996 (Nueno and Ghemawat, 2002). And when the Canadian manufacturing firm Blinds To Go set up a manufacturing plant in New Jersey, moving an experienced manager on site proved vital to improve the new plant's production efficiency (Menor and Mark, 2001).

Still, we know little if anything about the influence of geographic frictions on the optimal managerial organization of firms. Recent papers formalize the idea that adding a layer of middle managers allows firms to increase efficiency as they grow, and assemble

This chapter is based on joint work with Anna Gumpert and Manfred Antoni.

empirical evidence consistent with this hypothesis (e.g., Caliendo and Rossi-Hansberg, 2012; Caliendo et al., 2015a, b; Friedrich, 2016). However, existing studies focus on single establishment firms and disregard the possibility of multi-establishment production.

In this chapter, we show that geographic frictions between the headquarters and *one* subordinate establishment affect the optimal managerial organization of *all* establishments of a multi-establishment firm. Firm organization affects firm performance. Thus, our result implies that prior studies may underestimate the impact of geographic frictions on firm performance, because they focus on subordinate establishments. More generally, the main implication of this chapter is that analyses of the impact of local conditions at the establishment level provide only a partial picture of their total effect on multi-establishment firms.

We motivate our study by a set of stylized facts on multi-establishment firm organization. To derive the facts, we assemble a new linked firm-establishment-employee data set from administrative sources in Germany. Our data set is ideally suited to study multi-establishment firm organization because it combines detailed data about the employees of a firm and information about its geography. We summarize our findings in three facts.

First, multi-establishment firms prefer locations that are geographically close to their headquarters for their subordinate establishments. The location probability increases with the market potential of a location and decreases with the wages and the land prices relative to the headquarter location.

Second, multi-establishment firms are more hierarchical than single establishment firms. On average, multi-establishment firms have 2.0 management layers, whereas single establishment firms have 1.4 management layers. The difference is robust to controlling for firm characteristics. In particular, the difference persists conditional on firm size. The difference is related to geography: the number of managerial layers of multi-establishment firms increases with the distance of the subordinate establishments to the headquarters and with the area that the establishments cover.

Third, multi-establishment firms reorganize gradually. That is, they do not change the number of layers firm-wide, but add or drop layers establishment by establishment.

These facts suggest that geographic frictions affect both the location and the organization decisions of multi-establishment firms, and that the establishments are relevant units for the managerial organization of firms. We propose a model to explain why firms choose to organize their employees in multiple establishments, and why this decision affects the managerial organization. We consider a setting with two locations. Each firm consists of a CEO, production workers, and, optionally, one or more layers of middle managers. The CEO is located at the headquarters. The production workers and possible middle managers may be located at either or both locations. The CEO provides managerial services that are complementary to the labor input of the production workers in output production. The key assumption of the model is that the CEO is a resource of limited supply for the firm. He has only one unit of time. The location of the production workers determines the amount of time that the CEO needs to spend to provide managerial services. To release CEO time, the firm can hire middle managers that provide a subset of the managerial services. However, hiring middle managers entails quasi-fixed costs.

As point of reference, we first derive the optimal managerial organization if the CEO and the production workers are located in a single establishment. The CEO always fully uses his time to provide managerial services because they are complementary to the production workers' labor input. The larger the total output of the firm is, the more production workers it hires. The more production workers are to be managed, the more costly it is for the firm that only one unit of CEO time is available. The firm adds a layer of middle managers if the benefit of releasing CEO time outweighs the quasi-fixed costs of the middle managers (consistent with Caliendo and Rossi-Hansberg, 2012).

We next consider the multi-establishment case and allow the firm to hire employees at both locations. If the firm chooses to employ production workers at both locations, it optimally allocates the output such that the marginal production costs are equal across establishments, and the time of the CEO such that the marginal benefit of CEO time is equal across establishments. This insight is a key result of the model. The result implies that the managerial organization is interdependent across establishments. The establishment organization determines the marginal production costs and the marginal benefit of CEO time. As firm level optimization requires that these outcomes are equal across establishments, the organization decisions at the establishment level are interlinked. The interdependence affects the impact of firm size and geographic frictions on the organization.

Concerning firm size, the larger the total output is, the more production workers are hired, as in the single establishment case. The larger the total output is, the more costly the limit of CEO time therefore is for the firm. The firm can hire middle managers either at one or both establishments. Hiring middle managers at only one establishment entails lower quasi-fixed cost than hiring them at both establishments. The middle managers decrease the marginal benefit of CEO time at the establishment. They release CEO time that is reallocated to the other establishment to equalize the marginal benefit of CEO time across establishments. Middle managers at one establishment are thus beneficial for both establishments. Multi-establishment firms therefore add a layer of middle managers at one establishment at a lower firm size than single establishment firms. At the other establishment, they add a layer at a larger firm size than single establishment firms. This result arises because the middle managers are substitutes across establishments: The middle managers hired at the first establishment already release CEO time, thereby decreasing the need for middle managers at the second establishment.

Concerning geographic frictions between the subordinate establishment and the headquarters, they affect the organization and location decisions of firms. The frictions increase the amount of CEO time needed to provide managerial services and thus the costs of the CEO time limit. In response, the firm adjusts the organization. Importantly, it adjusts the organization of both establishments to maintain that the marginal benefit of CEO time and the marginal production costs are equal across establishments. The more costly the limit of CEO time is for the firm, the more beneficial it is for the firm to hire middle managers. Higher geographic frictions thus increase the number of management layers of a multi-establishment firm. The middle managers and other organizational adjustments mitigate, but do not reverse the positive impact of geographic frictions on firms' production costs. The firm therefore only produces at both locations if lower wages or advantages such as avoiding transport costs between locations outweigh the higher costs of providing management services for the CEO.

The model reproduces the facts documented in the data. Multi-establishment firms have more management layers than single establishment firms and reorganize gradually, establishment by establishment. Geographic frictions increase the number of layers and decrease the appeal of multi-establishment production.

The key implication of the model is that geographic frictions between the headquarters and one establishment have repercussions on the organization of all establishments of the firm. In the final step of our study, we exploit the opening of high-speed train routes in Germany during our sample period to provide evidence for this prediction. The train routes affect the travel time between subordinate establishments and the headquarters and thus provide plausibly exogenous variation of the costs of managing subordinate establishments from the headquarters. The new train connections provide the fastest mode of travel between locations: they are faster than cars or planes (if one accounts for waiting times at the airport). We study their impact using a differences-in-differences econometric strategy. We run regressions of outcomes at the firm level, the level of the treated subordinate establishment and of untreated subordinate establishments. With view to the model predictions, we exclude the "untreated" establishments of "treated" multi-establishment firms from the control group for the "treated" establishments. We find that firms benefiting from faster travel times grow faster than other firms. They reallocate employment to the establishment that is faster to reach. Importantly, we find that the new train routes affect both the organization of the "treated" and the "untreated" subordinate establishments of the multi-establishment firm. This is consistent with the interdependence of establishment organization implied by the model.

The key insight of the model is that geographic frictions not only affect a specific subordinate establishment, but all establishments of a multi-establishment firm. This insight is particularly relevant for the literature on multi-establishment and multinational firms. In the literature on multi-establishment firms, the determinants of firm performance receive increasing attention. Recent papers uncover that distance to the headquarters and other geographic frictions decrease investment, productivity and longevity of subordinate establishments of multi-establishment firms (e.g., Giroud, 2013; Kalnins and Lafontaine, 2013).[1] We show that the impact of geographic frictions exceeds their effect on the specific subordinate establishments. In the literature on multinational firms, headquarter

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

inputs are typically considered public goods within the firm (e.g., Helpman et al., 2004; Antràs and Yeaple, 2014, for a survey). We show that the public good assumption may apply to patents or trademarks, but not to managerial inputs. Geographic frictions and other local conditions thus affect not only the local foreign affiliate, but the network of a multinational firm.

Beyond the literature on multi-establishment and multinational firms, our study contributes to the literature on firm organization and management by showing that geographic frictions are a determinant of firm organization. To develop our model, we build on the literature of firms as knowledge hierarchies (for an overview, see Garicano and Rossi-Hansberg, 2015). A series of papers formalizes the idea that firms add management layers as they grow to maintain their productivity, and provides empirical evidence for it (e.g., Caliendo and Rossi-Hansberg, 2012; Caliendo et al., 2015a, b; Friedrich, 2016). Similar theoretical predictions result from a monitoring hierarchy framework (e.g., Chen, 2017; Chen and Suen, 2017).[2] The literature focuses on size as main determinant of organization. Geographic frictions have been largely neglected, even though multi-establishment firms are among the largest firms in developed economies and account for a substantial share of aggregate employment.[3] While we implement the model in the knowledge hierarchy framework, we stress that our main results do not depend on this specific framework and would hold in a monitoring framework.

To the best of our knowledge, Charnoz et al. (2015) is the only study of the impact of geographic frictions on firm organization. This empirical paper shows that high speed train routes decrease the share of managers at subordinate establishments and increase establishment performance. This chapter combines theoretical and empirical analyses. This allows us to provide a novel and nuanced interpretation of the regression results on the impact of high speed train routes. Further, based on the insights from the model, we take the impact of lower travel times on "untreated" subordinate establishments into account.

[2] In the broader literature on the hierarchical organization of firms, Rajan and Wulf (2006) and Guadalupe and Wulf (2010) study the organization of management positions in 300 large publicly traded U.S. firms.

[3] Gumpert (2018) contains a knowledge hierarchy model where firms produce at more than one location, but with a fixed number of layers. Crèmer et al. (2007) study firm language in a setting with multiple divisions.

This chapter also offers a novel perspective on the recent management literature. Bloom et al. (2017) document that half of the total variation in management practices between different U.S. establishments is due to variation between establishments within the same firm. They argue that larger firms may find it harder to align management practices across establishments (p. 10). Our model implies that heterogeneous management practices in multi-establishment firms may reflect asymmetries in the optimal organization of employees across establishments. Implementing managerial practices requires managerial time. Asymmetries in the number of managerial layers and the amount of CEO time allocated to an establishment may manifest in heterogeneous managerial practices.

The chapter is structured as follows. The next section describes the data. Section 3.3 presents the facts on multi-establishment firm organization. Section 3.4 develops a model of firm organization consistent with the facts. Section 3.5 presents the evidence from the opening of high speed train routes. The last section concludes.

3.2 Data

Our study requires information both on the geographic location of the establishment(s) and the managerial organization of firms.

3.2.1 Data construction and descriptive statistics

We construct a detailed linked firm-establishment-employee data set for Germany that is uniquely suited to study multi-establishment firms.[4] The data contain information on the legal form and sales of firms, and the location at the county level, three digit sector, and age of each establishment. We observe all employees of the establishments subject to social security contributions on 30 June every year. For each employee, the data include the occupation, age, gender, level of education, employment history and wages. The data cover the period 1998-2014. Each employee, establishment and firm has a unique identifier that allows following the units of observation over time.

[4]We use the same dataset as in chapter 2. To make each chapter self-contained, we describe the data for each chapter separately.

We assemble the data set from two sources. The universe of Social Security records provides the data on employees and establishments. The Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) makes the data available for research. The employee history contains the data on the employees. The Establishment History Panel and the extension file entry and exit contain the information on the establishments.[5] The ORBIS database of Bureau van Dijk (BvD) contains balance sheet information of firms. We combine the Social Security records and the ORBIS database using record linkage techniques. The algorithm exploits the regulation that the establishment names in the Social Security data have to contain the firm name. We identify the headquarters (HQ) establishment of a firm as the establishment with the same zip code or locality as the firm.[6] Appendix C.1.1 contains details on the components of our data set and the record linkage procedure.

The data set is an unbalanced panel. We use the 2012 cross section for cross-sectional analyses, because it contains the maximum number of establishments. The panel analyses use the period 1998-2010. We exclude the year 2011 because of changes in the occupational classification in that year (for details, see Appendix C.1.2). Consistent with the literature, we restrict our sample to full-time employees (e.g., Card et al., 2013; Dustmann et al., 2009). We focus on firms with at least 10 employees in all years. 99% of the firms dropped due to this requirement are small single establishment firms.

Table 3.1 provides descriptive statistics of the 2012 cross-section. As the upper panel shows, our sample comprises 109 thousand firms that consist of 144 thousand establishments and employ 6.4 million individuals. The data cover almost one third of total full time employment subject to social security contributions in Germany in 2012 (Bundesagentur für Arbeit, 2016).[7] We do not observe sales for all firms, but only the larger firms due to limitations of the BvD data. Though only 9 percent of all firms in our sample are multi-establishment firms, 31 percent of establishments belong to and 34 percent of employees work for them. The sample covers all sectors. The share of multi-

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

[6]The Social Security records contain the address of each establishment and the ORBIS database contains the address of the firm. We are allowed to use the address to identify headquarters, but are allowed to use only the county of the establishments in our analyses for confidentiality.

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

Table 3.1: Descriptive statistics, SE vs. ME firms, 2012 cross section

Units of observation	N	% share ME firms						
Firms	109,348	9.0						
with non-missing sales	54,035	9.4						
Establishments	144,428	31.0						
Employees	6,355,914	34.0						
Descriptive statistics	**N**	**ME**	**Mean**	**SD**	**p25**	**p50**	**p75**	**p95**
# employees per firm	99,524	0	42	92	13	21	39	133
	9,824	1	222	1979	22	50	127	650
Sales per firm (M €)	48,976	0	29	750	2	4	9	73
	5,059	1	350	4,238	4	17	79	573

Table 3.2: Descriptive statistics, ME firms, 2012 cross section

Descriptive statistics	N	Mean	SD	p50	p75	p95
# establishments per ME firm	9,824	4.6	19.6	2	3	10
# sectors per ME firm	9,824	1.6	0.9	1	2	3
# employees per establishment	44,904	48	430	8	24	156
Maximum distance to HQ in km	9,824	218	189	39	167	546
Minimum area covered in km^2	3,584	$30,075$	$41,712$	$6,933$	$49,717$	$124,564$

establishment firms is similar across sectors. It ranges from 4.5 percent in construction to 7.5 percent in manufacturing, the largest broad sectoral category, and a maximum of 12 percent in retail and services.

As the descriptive statistics in the lower panel show, multi-establishment (ME) firms are substantially larger than single establishment (SE) firms in terms of their employees 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 fivefold. Sales of multi-establishment firms are fourfold those of single establishment firms at the median.

Table 3.2 documents the heterogeneity among the group of multi-establishment firms. While more than half of multi-establishment firms have two establishments, the largest five percent have ten or more establishments. Most multi-establishment firms are active in only one three-digit sector. Even at the higher end of the distribution, the number of sectors is significantly lower than the number of establishments. The size of the establishments varies with a larger standard deviation than the one for single establishment firms, which results because the size cut-off is not binding at the establishment level for multi-establishment firms. To capture firm geography, we use the distance in kilometers between a subordinate establishment and the headquarters and the minimum area in square kilometers covered by all establishments. The latter only applies to firms with at least two subordinate stablishments. Half of all multi-establishment firms do not have establishments that are farther than 39 km from their headquarters. At the top of the distribution, the distance exceeds 540 km, which is about two thirds of the maximum possible distance within Germany. The distribution of the area is similarly skewed.

3.2.2 Measures for the managerial organization

We use the occupation of the employees to construct three measures of the managerial organization of firms. First, we count the number of managerial layers of firms. We assign employees to four layers depending on their occupation (as Caliendo et al., 2015b):

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

We transfer the mapping in Caliendo et al. (2015b) based on the French classification of occupations to the German classification using official correspondence tables (Friedrich, 2016, uses an analogous procedure for Danish data). We treat the layer at the lowest level in each establishment as non-managerial. We count the number of layers above the lowest layer per firm. The lowest layer contains employees at level 0 in 98 percent of firms. Multi-establishment firms may separate management and production, which is why we cross-check our findings treating the lowest level in the firm as non-managerial. Appendix C.1.3 provides details on our procedure and a list of occupations by level.

The two other measures are shares of managerial occupations in the wage sum, where we determine which employees have managerial occupations in two ways. On the one hand, we build on the assignment of employees to managerial layers and treat all employees above the lowest level as managerial. The establishments report the occupations of the employees in the social security data. In multi-establishment firms, establishments may assign different occupations to similar employees. Cross-checking the results on the number of layers with the management share helps ensure that our results are robust to this possibility. On the other hand, we use the assignment of Blossfeld (1983, 1987, see Appendix C.1.3 for the list of managerial occupations). The assignment builds on research from sociology and is part of establishment history panel. Managers are employees in occupations that have control or decision-making power over the use of production factors as well as high-level officials in organizations (Blossfeld, 1983, p. 208).

One may be concerned to which extent the occupation classification captures the managerial position of employees in a meaningful way. Using survey data, we show that the tasks and job characteristics of occupations are systematically different between layers in ways that plausibly reflect different roles of employees within firms (see Appendix C.1.4).

3.3 Facts on firm location and organization

This section describes the location and organization patterns of multi-establishment firms. We first describe how geographic frictions between a location and the headquarters affect the decision of multi-establishment firms where to locate an establishment as well as establishment size. Taking the location decisions as given, we then describe the managerial organization of multi-establishment firms in the cross-section and over time and compare it to the organization of single establishment firms.

3.3.1 Distance to headquarters decreases location probability

Table 3.3 describes the location patterns of multi-establishment firms. Columns 1 to 4 present the results of probit regressions that relate y_{ij}, a dummy variable that is equal to one if firm i maintains a subordinate establishment in county j and firm and county characteristics:

$$\Pr(y_{ij} = 1) = \Phi(\beta_0 + \beta_1 \mathbf{x}_i + \beta_2 \mathbf{x}_j). \tag{3.1}$$

Consistent with a negative impact of geographic frictions between the headquarters and a subordinate establishment on establishment performance, firms are the less likely to locate an establishment in a county, the more distant the county is from the headquarters. A larger market potential relates positively to the location probability, indicating market-seeking motives. Lower wages and land prices in the county relative to the headquarters are also positively related to the location probability, which points to cost-cutting motives. Finally, larger multi-establishment firms are more likely to set up subordinate establishments.

Columns 5 and 6 present the results of OLS regressions that relate the number of employees of a subordinate establishment to county characteristics. The regressions control for firm fixed effects to account for the possibility that larger firms have more establishments. The sample therefore only includes multi-establishment firms with at least two subordinate establishments. Subordinate establishment size is negatively related to the distance between a county and the headquarters, again consistent with a negative impact of geographic frictions between the headquarters and the subordinate establishment on establishment performance. Larger market potential relates positively and higher wages relate negatively to establishment size. Unlike higher wages, higher land prices are positively related to establishment size. A possible explanation for the different sign patterns is that land is a fixed cost for production. Thus, it is worthwhile to maintain only larger establishments at higher land price locations.

Fact 1 summarizes our findings:

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

Table 3.3: Location probability and establishment size, ME firms, 2012 cross section

Dependent variable	Location probability, Probit				Log est. emp., OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Log distance to HQ	-0.263^{***}	-0.247^{***}	-0.309^{***}	-0.259^{***}	-0.076^{***}	-0.080^{***}
	(0.018)	(0.019)	(0.026)	(0.027)	(0.017)	(0.018)
Log market potential	0.543^{***}	0.635^{***}	0.678^{***}	0.692^{***}	0.372^{***}	0.358^{***}
	(0.029)	(0.032)	(0.037)	(0.035)	(0.037)	(0.041)
Relative wages	-0.567^{***}	-0.552^{***}	-0.481^{***}	-0.590^{***}	-0.112	-0.212^{**}
	(0.168)	(0.163)	(0.114)	(0.145)	(0.080)	(0.079)
Relative land prices		-0.046^{***}	-0.042^{***}	-0.044^{***}		0.022^{***}
		(0.007)	(0.006)	(0.006)		(0.006)
Log # employees			0.252^{***}			
			(0.014)			
Log sales				0.149^{***}		
				(0.011)		
# of observations	3,934,612	3,415,095	3,225,429	1,757,916	21,496	19,203
# of firms	9,812	9,255	8,741	4,758	3,066	2,773
HQ sector dummies	N	N	Y	Y	N	N
HQ county dummies	N	N	Y	Y	N	N
Legal form dummies	N	N	Y	Y	N	N
Firm fixed effects	N	N	N	N	Y	Y

Table presents coefficient estimates. Standard errors clustered at HQ-county-level in parentheses. ** $p < 0.01$, *** $p < 0.001$. The sample includes multi-establishment firms with at least 10 employees. *Dependent variable:* (1)-(4): dummy variable that is equal to one if firm i has a subordinate establishment in county j (HQ counties are excluded), (5)-(6): log number of employees at subordinate establishment. *Independent variables: Log distance to HQ:* log distance between county j and county of HQ of firm i in km; *Log market potential:* distance weighted average of the GDP of county j and surrounding counties; *Relative wages/land prices:* wages/land prices in county j relative to wages in county of headquarters of firm i; *Log number of employees:* log number of employees of firm i; *Log sales:* log sales of firm i. Wages are calculated as average wages in a county excluding the respective firm. Number of observations varies because of covariate availability. Distance, market potential and relative land prices are computed based on data on the coordinates of municipalities as well as GDP and land prices of counties provided by the German Statistical Office. Relative wages are from the German Social Security data.

3.3.2 Distance to headquarters increases managerial share

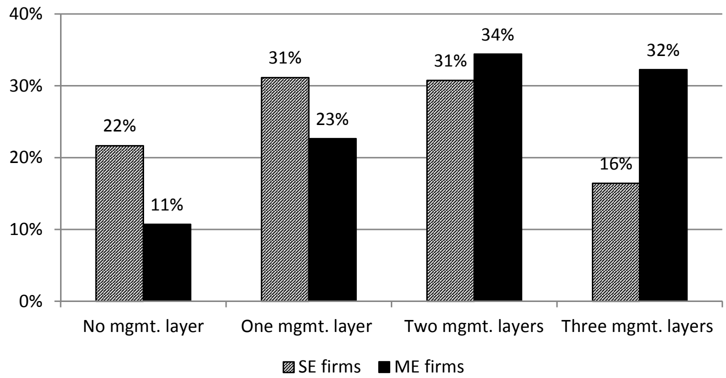
We proceed in two steps to describe the relation of geographic frictions and firm organization. First, we compare the managerial organization of single and multi-establishment firms. This helps understand whether the number of establishments affects the managerial organization. Second, we restrict the sample to multi-establishment firms and relate geographic frictions and the managerial organization.

Figure 3.1 plots the number of management layers by firm type. On average, firms have 1.46 management layers with a standard deviation of around 1 in the 2012 cross-section. Multi-establishment firms are more hierarchical than single establishment firms: the average number of management layers in multi-establishment firms is 1.88 and higher than the average number of 1.42 in single establishment firms. The distribution is dissimilar between single and multi-establishment firms. The distribution has an inverse U-shape for single establishment firms. A third have one or two management layers respectively. 22 percent have no management layer, and less than one fifth have three management layers. In contrast, two thirds of multi-establishment firms have two or three management layers. Only around ten percent do not have a management layer and 23 percent have one management layer.[8] Likewise, the managerial share of multi-establishment firms is larger than the one of single establishment firms. When we define the share by layer, employees in managerial occupations in multi-establishment firms command 34% of the wage sum—six percentage points more than in single establishment firms. According to Blossfeld (1987)'s definition, the share of managerial employees in total wages in multi-establishment is 1.5 times the share in single establishment firms (9% vs. 6%).

While the differences in Figure 3.1 may be driven by a firm's number of establishments, they may likewise result from the differences in size between single and multi-establishment firms documented in Table 3.1. Table 3.4 presents the results of regressions that condition on size as determinant of the number of management layers and

[8]Some firms do not have management layers for two reasons. First, social security data only contain information on employees that pay social security contributions. Owner-managers are thus only included if they pay themselves a wage. Our results are robust to separate estimation by legal form (see Appendix Table C.8). Second, the data contain only one occupation per employee. Managers of small firms may be attributed a production occupation if they execute such an occupation for much of their time.

Figure 3.1: Number of management layers by firm type, 2012 cross-section.



The figure plots the distribution of the number of management layers separately for SE and ME firms in the 2012 cross-section. The sample includes firms with at least ten employees and non-missing legal form. 82% of firms have consecutive layers. Appendix Table C.3 displays the share of firms with consecutive layers by firm type and number of layers.

take differences in the sector, legal form and location of firms into account. Specifically, columns 1 to 4 estimate

$$\#\,\text{management layers}_i = \exp\left(\beta_0 + \beta_1\, D_{\text{ME firm},i} + \beta_2\,\text{size}_i + \alpha_l + \alpha_n + \alpha_s\right), \qquad (3.2)$$

where i refers to the firm, l to its legal form, n to the county of the headquarters, s to the headquarter sector, $D_{\text{ME firm},i}$ is a dummy equal to one for ME firms and zero otherwise, and α denotes fixed effects. As the mean and variance of the number of management layers are approximately equal, the Poisson model is a reasonable approximation of the data. Columns 5 and 6 use the share of managerial occupations in the wage sum as dependent variable and relate it to the control variables using OLS regressions. We do not condition on size because the total wage sum, the denominator of the managerial share, is strongly correlated with the size measures.

Through columns 1 to 4, multi-establishment firms have a significantly higher number of layers than single establishment firms. The coefficients in column 1 imply that multi-establishment firms have 9 percent more management layers than single establishment firms. The effect is equivalent to increasing the number of non-managerial employees by 50 percent. As column 2 shows, the multi-establishment firm dummy does not reflect a non-linear size effect. The effect is smaller, but still positive and significant when we control for sales in columns 3 and 4. The decreases in effect sizes is likely partly due

Table 3.4: Regression results, managerial organization, 2012 cross-section

Dependent variable	# mgmt. layers, Poisson				Managerial share, OLS	
					Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{\text{ME firm}}$	0.085***	0.058***	0.021***	0.023***	2.024***	1.238***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.231)	(0.122)
Log # non-mg. employees	0.148***	−0.050***		−0.004		
	(0.002)	(0.011)		(0.003)		
Log # non-mg. employees^2		0.030***				
		(0.002)				
Log sales			0.182***	0.184***		
			(0.002)	(0.003)		
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
# observations	105,949	105,949	53,566	53,566	105,947	105,947

2012 cross-section, only firms with at least 10 employees. Robust standard errors in parentheses. *** p < 0.001. *Dependent variable:* 1-4 number of management layers, 5 managerial share in wage sum, layer definition, 6 managerial share in wage sum, Blossfeld. *Independent variables:* $D_{ME\ firm}$: 1 if firm is ME firm, 0 otherwise; *Log # non-mg. employees:* log number of employees at lowest layer of establishments; *Log sales:* log sales of the firm. Constant included.

to the non-random availability of the sales data. The ORBIS contains sales information only for the larger firms in the sample. Columns 5 and 6 show similar results for the managerial share. If defined based on the managerial layers, the managerial share in multi-establishment firms is two percentage points higher than the share in single establishment firms. At the mean, this difference is equivalent to an increase by seven percent. According to Blossfeld (1987)'s definition, the difference is 1.2 percentage points. At the mean, this difference is equivalent to a 20 percent increase.

To explore whether the higher number of management layers of multi-establishment firms is related to geographic frictions, we restrict the sample to multi-establishment firms and re-estimate equation (3.2) taking into account geography:

$$\#\,\text{management layers}_i = \exp\left(\beta_0 + \beta_1 \text{geography}_i + \beta_2 \text{size}_i + \alpha_l + \alpha_n + \alpha_s\right), \qquad (3.3)$$

where i now refers to the multi-establishment firm. We estimate analogous OLS regressions using the managerial share as dependent variable. We employ two measures of firm

Table 3.5: Regression results, managerial organization of ME firms, 2012 cross-section

Dependent variable	# mgmt. layers, Poisson				Managerial share, OLS			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log distance to HQ	0.019*** (0.004)		0.010*** (0.004)		0.474* (0.202)		0.233* (0.096)	
Log area spanned by firm		0.011*** (0.003)		0.011*** (0.003)		0.195 (0.143)		0.181** (0.070)
Log # non-mg. employees	0.139*** (0.004)	0.117*** (0.006)						
Log sales			0.125*** (0.004)	0.100*** (0.005)				
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
# firms	9,275	5,033	3,320	1,984	9,275	3,320	9,275	3,320

2012 cross-section, only multi-establishment firms with at least 10 employees. Columns 2, 4, 6, 8 include only ME firms with at least two subordinate establishments. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, 5-6 managerial share in wage sum, layer definition, 7-8 managerial share in wage sum, Blossfeld. *Independent variables: Maximum log distance to headquarters*: log of maximum distance between subordinate establishment and headquarters in km; *Log area spanned by firm*: log of minimum area covered by establishments in square kilometers; others see Table 3.4.

geography, the maximum log distance in kilometers between a subordinate establishment and the headquarters, and the minimum area in square kilometers covered by all establishments. Distance is defined for all multi-establishment firms, whereas the area is only defined for firms with at least two subordinate establishments. We take the maximum distance of the subordinate establishments to the headquarters if the firm has more than one subordinate establishment; using the mean distance yields similar results. Firm size controls both for the positive effect of size on the number of layers and for the possibility of larger firms investing at farther destinations.

Table 3.5 presents the regression results for the 2012 cross-section. The regression results show that both distance and area have a positive impact on the number of management layers in a firm. According to column 1, doubling the maximum distance of an establishment to the headquarters is associated with a 2 percent increase of the number of layers. The magnitude of the effect is about a sixth of the elasticity of the number of layers with respect to the number of non-managerial employees. The effect is robust to using sales as size measure in column 3. The impact of the log area in columns 2 and 4 is similar. Likewise, the managerial share is positively related both to the maximum distance of the establishments and the area they span. Coefficients are significant, except for column 6 (P-value 17%).

Fact 2 summarizes our findings:

Fact 2. *Multi-establishment firms have more management layers than single establishment firms of the same size, legal form, sector and headquarter county. The number of management layers of multi-establishment firms relates positively to the distance between the headquarters and the subordinate establishments and the area spanned by the establishments, conditional on firm characteristics. The same holds for the managerial share.*

Robustness. Appendix section C.2 shows that our results are robust to a series of checks. First, multi-establishment firms may separate management and production geographically. We therefore replicate our analysis treating the lowest-level layer in each firm as non-managerial layer. Appendix Table C.4 shows that the differences between single and multi-establishment firms are even stronger with this definition of management layers. According to the estimates in column 1, being a multi-establishment firm quantitatively relates to the number of management layers as doubling the number of non-managerial employees. The coefficients in columns 2-4 are also larger than the corresponding effects in Table 3.4. The coefficient for the managerial share is twice its counterpart in column 5 of Table 3.4. Appendix Table C.5 shows that the relation of geographic frictions and the number of management layers and the managerial share defined at the firm level is similar as in the baseline regressions.

Second, Appendix Tables C.6 and C.7 replicate the cross-section regression results for the 1998-2010 panel. Third, we explore potential sources of omitted variables bias. The legal form affects whether owner managers are subject to social security contributions. Appendix Tables C.8 splits the sample by legal form to allow for heterogeneity in the coefficients across legal form categories. The estimated coefficient of the multi-establishment firm dummy is robust across legal form groups. In Table C.9, we exclude multi-establishment firms with establishments in different sectors, or subordinate establishments in the headquarter county, as well as large multi-establishment firms with more employees than the 95th percentile of single establishment firms. The estimated coefficients are very similar to the baseline coefficients. Finally, the specification in Table 3.5 imposes a linear relation of distance or area and firm organization. Table C.10 uses quartile dummies. The table shows that the coefficients in Table 3.5 are driven by the top quartile.

3.3.3 Multi-establishment firms reorganize gradually

Facts 1 and 2 show that the location and organization of multi-establishment firms are related to geographic frictions in the cross-section. So far, the analysis refers to the managerial structure at the *firm* level. For multi-establishment firms, heterogeneity of the organization of the *establishments* may, however, be important. In fact, the managerial organization of subordinate establishments is rarely a copy of the one in the headquarters: in around 50 percent of all multi-establishment firms, the number of managerial layers at the headquarters exceeds the number of layers at all subordinate establishments. Even if the number of layers is similar, the level of management often differs.

To understand whether the heterogeneity among the organization of the establishments is a constant feature of multi-establishment organization, we study changes in the managerial organization over time. Table 3.6 displays the percentage shares of firms that transition from a number of managerial layers in year t to a possibly different number of managerial layers in year $t + 1$ separately for single and multi-establishment firms. The propensity of firms to reorganize is similar across the two groups. The managerial organization is sluggish: at least four fifth of firms in both groups keep their number of managerial layers across periods. In case that firms change the number of layers, they typically add or drop one layer. In only one instance, one percent of firms add or drop more than one layer.

Table 3.7 digs deeper into the organization of multi-establishment firms and considers their dynamics at the establishment level. To summarize the managerial organization of multi-establishment firms with a possibly different number of establishments, the table counts the number of managerial layers at the headquarters and the maximum number of managerial layers at the subordinate establishment.

Table 3.6: Transition dynamics of the managerial organization, by firm type

# layers	SE firms					ME firms					
in t/t+1	0	1	2	3	# firms	0	1	2	3	SE	# firms
0	**92**	7			169,766	**85**	9			6	11,714
1	6	**87**	7		213,855	5	**83**	7		5	22,480
2		10	**83**	6	142,753		8	**81**	7	5	21,019
3		1	10	**89**	82,092			6	**91**	4	23,015

The table displays, separately for SE and ME firms, the percentage share of firms that reorganize from a number of managerial layers in year t (given in the rows) to a possibly different number of managerial layers in year $t+1$ (given in the columns). Cells that contain fewer than 1% of observations are left empty to ease readability. Sample: 1998-2010 panel of firms with at least 10 employees in all years. Fewer than 1% of firms exit. Diagonal in bold.

Table 3.7: Transition dynamics of the managerial organization within ME firms

# layers in $t/t+1$	0/0	1/<1	1/1	2/<2	2/2	3/<3	3/3	SE	# firms
HQ 0/ sub.e. 0	**85**	5						6	11,714
HQ 1/ sub.e. 0	6	**75**	4	6				8	10,284
HQ 1/ sub.e. 1	1	6	**75**	7				2	7,865
HQ 2/ sub.e. 0,1		4	4	**77**	2	6		6	13,619
HQ 2/ sub.e. 2			1	10	**70**	9	1	1	3,727
HQ 3/ sub.e. 0,1,2				5	2	**85**	3	5	15,249
HQ 3/ sub.e. 3						8	**88**	1	5,323

The table displays the percentage share of firms that change from a managerial structure in year t (given in the rows) to a managerial structure in year $t+1$ (given in the columns). The figure in front of the slash denotes the number of managerial layers of the headquarters. The figure behind the slash denotes the maximum number of managerial layers at the subordinate establishments. Cells that contain fewer than 1% of observations are left empty to ease readability. Sample: 1998-2010 panel of firms with at least 10 employees in all years. Firms with a higher number of layers at the subordinate establishment than at the HQ dropped for readability. Fewer than 1% of firms exit. Diagonal in bold.

Two findings are notable. First, the managerial organization at the establishment level is less stable than the managerial organization at the firm level: there is less mass on the diagonal of Table 3.7 than on the diagonal of the right panel of Table 3.6. Second, multi-establishment firms reorganize gradually and add or drop layers at one establishment at a time. For example, among multi-establishment firms with two layers both at the headquarters and the subordinate establishments, 9 percent add a layer at the headquarters and 10 percent drop a layer at the subordinate establishments, but only 1 percent of firms choose a lower or higher number of layers across all establishments. The latter adjustment does not show up as reorganization at the firm level.

We interpret these findings as evidence that the establishments are important entities for the organization of multi-establishment firms. Organizational adjustments do not only take place at firm level, but at establishment level. Fact 3 summarizes our finding.

Fact 3. *Multi-establishment firms reorganize gradually and add or drop layers at one establishment at a time.*

3.4 A model of firm organization with multiple establishments

To explain why the number of establishments of a firm affects the managerial organization, we develop a model of the organization of employees in multiple establishments. We allow firms to endogenously choose both the number of establishments and the managerial organization. The key assumption in our model is that CEO competency is a resource of limited supply for a firm, because the CEO has only one unit of time. We solve the optimization problem in three steps. In section 3.4.2, we derive the optimal managerial organization of a single establishment firm. In section 3.4.3, we derive the optimal organization of a multi-establishment firm. We first consider a setting with wage differences as only motive for having two establishments and then study a setting with both wage differences and market access considerations.

3.4.1 Set-up

We consider an economy with two locations, $j = \{0, 1\}$. The N_j agents per location each supply one unit of time to the labor market. The agents are immobile, so local wages w_j can differ. We choose indexes such that $w_0 \geq w_1$. The agents derive utility from consuming differentiated products i:

$$U(x(\alpha_i)) = \left(\int_A \alpha_i^{\frac{1}{\sigma}} x(\alpha_i)^{\frac{\sigma-1}{\sigma}} M dG(\alpha) \right)^{\frac{\sigma}{\sigma-1}}. \tag{3.4}$$

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

To simplify the exposition, sections 3.4.2 and 3.4.3 analyze the problem of finding the optimal organization of a firm with taste draw α in location 0. Section 3.4.4 studies the competition among many firms i in the goods market.

Production. Production is a problem solving process based on labor and knowledge (as in Caliendo and Rossi-Hansberg, 2012; Garicano, 2000). Every unit of labor 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 can be solved if it is realized within the knowledge interval. The problems follow a distribution with the exponential density $f(z) = \lambda e^{-\lambda z}$, where $z \in [0, \infty)$ refers to the domain of possible problems and λ denotes the predictability of the production process. Combining n units of labor and knowledge $\bar{z}$ yields

$$q = n\left(1 - e^{-\lambda \bar{z}}\right)$$

units of output, where $1 - e^{-\lambda \bar{z}}$ is the value of the cumulative distribution function.

The firm hires agents on the labor market to put labor and knowledge in production. The firm's employees put in labor by spending their time generating problems. To use knowledge in production, the employees have to learn it first. 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 with each other, so they can leverage differences in their knowledge. Communication is costly: an employee in location j spends θ_{kj} units of time listening to problems communicated by an employee in location k. Communication across space is more costly than communication within a location: $1 > \theta_{10} \geq \theta_{00} > 0$. The communication costs are symmetric: $\theta_{10} = \theta_{01}, \theta_{11} = \theta_{00}$. If an employee does not know how to solve a problem, he cannot tell who knows, but has to find a competent fellow employee.

Organization. The firm organizes its 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 put in labor and some knowledge in production. They generate problems and solve those problems that are realized in their knowledge interval. We call the employees at the higher layers $\ell \geq 1$ managers. They put only knowledge in the production process and spend their time listening to unsolved problems from the employees at the next lower layer. The highest managerial layer consists of the CEO. We assume that each firm has exactly one CEO. The knowledge levels of the employees are overlapping, so employees at layer ℓ know the knowledge of employees at layer $\ell - 1$ and more.[9] 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. As the density of problems is decreasing in z, the knowledge of the production workers covers the solution to the most common problems, whereas higher layers also know the solutions to problems that occur more rarely. This minimizes the probability that costly communication is necessary.[10]

The communication costs θ_{jk}, the learning costs c, the predictability of the production process λ and the taste α 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 taken as given. We take total output $\tilde{q}$ as given in sections 3.4.2 and 3.4.3, but endogenize it in section 3.4.4.

3.4.2 The optimal organization of single establishment firms

We first determine the optimal organization of a single establishment firm. The optimal organization minimizes the production costs. It consists of the number of managerial layers L, the number $n^{\ell}_{0,L}$ and knowledge level $z^{\ell}_{0,L}$ of employees per layer $\ell = 0, ..., L-1$, and the knowledge of the CEO $\bar{z}_{0,L}$. The indexes $0, L$ refer to the location of the firm $j = 0$

[9] We assume that knowledge levels are overlapping to simplify the optimization problem of the multi-establishment firm. With non-overlapping knowledge levels, in a multi-establishment firm, both overlaps and gaps between CEO and establishment knowledge may occur. This complicates the analysis without adding insights.

[10] Garicano (2000) shows that an optimal knowledge hierarchy features specialization and organization by frequency, i.e. only the lowest layer inputs labor and the knowledge of higher layers covers the rarer problems.

and the number of managerial layers L, because these variables affect the values of the other choices.

The optimal number of layers is given by

$$C(\tilde{q}) = \min_{L \geq 1} \tilde{C}_{0,L}(\tilde{q}). \tag{3.5}$$

The optimal number and knowledge levels of employees at all layers solve:

$$C_{0,L}(\tilde{q}) = \min_{\{n_{0,L}^{\ell}, z_{0,L}^{\ell}\}_{\ell=0}^{L-1}, \bar{z}_{0,L} \geq 0} \sum_{\ell=0}^{L-1} n_{0,L}^{\ell} w_0 \left(1 + c z_{0,L}^{\ell}\right) + w_0 \left(1 + c \bar{z}_{0,L}\right) \tag{3.6}$$

$$\text{s.t.} \quad n_{0,L}^{0} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right) \geq \tilde{q} \tag{3.7}$$

$$1 \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{L-1}} \tag{3.8}$$

$$n_{0,L}^{\ell} \geq n_{0,L}^{0} \theta_{00} e^{-\lambda z_{0,L}^{\ell-1}} \quad \forall \ell = 1, ..., L-1 \tag{3.9}$$

$$\bar{z}_{0,L} \geq z_{0,L}^{L-1}, \quad z_{0,L}^{\ell} \geq z_{0,L}^{\ell-1} \quad \forall \ell = 1, ..., L-1 \tag{3.10}$$

The production costs consist of the personnel costs for the employees and the CEO. Constraint (3.7) implies that the number of production workers and CEO knowledge have to suffice to produce total output $\tilde{q}$. According to constraints (3.8) and (3.9), the amount of time of the CEO and the managers limit the number of problems that can be communicated to them. The number of problems communicated to a higher layer is computed as the number of problems, $n_{0,L}^{0}$, multiplied with the communication costs, θ_{00}, and the probability that the problem is not yet solved, $e^{-\lambda z_{0,L}^{\ell-1}}$. Finally, knowledge levels are overlapping and positive (constraint 3.10).

Appendix C.3.1 contains the Lagrangian equation and the first order conditions. Two multipliers from the Lagrangian equation are key to characterizing the optimal organization. The multiplier for constraint (3.7), $\xi_{0,L}$, denotes the marginal production costs. The multiplier for constraint (3.8), $\varphi_{0,L}$, denotes the marginal benefit of CEO time. CEO time is fixed: only one unit is available. $\varphi_{0,L}$ reflects how costly this constraint is for the firm.

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

$$w_0 c = \frac{\lambda e^{-\lambda \bar{z}_{0,L}}}{1 - e^{-\lambda \bar{z}_{0,L}}} \xi_{0,L} \tilde{q} \tag{3.11}$$

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

The binding constraint (3.7) determines the number of production workers $n^0_{0,L}$ as a function of CEO knowledge. Constraint (3.8) 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 such that only the one unit of CEO time is used. The first order conditions imply that the knowledge levels of the production workers and managers at lower layers are a recursive function of the knowledge level at the highest layer:

$$e^{\lambda\left(z^{\ell-1}_{0,L} - z^{\ell-2}_{0,L}\right)} = \left(1 + c z^{\ell}_{0,L}\right) \frac{\lambda}{c} \quad \forall \ell = 2, ..., L-1, \tag{3.12}$$

$$e^{\lambda z^0_{0,L}} = \left(1 + c z^1_{0,L}\right) \frac{\lambda \theta_{00}}{c}. \tag{3.13}$$

Constraint (3.9) determines the number of middle managers as a function of the number of production workers and knowledge levels. Finally, the marginal production costs $\xi_{0,L}$ and the marginal benefit of CEO time $\varphi_{0,L}$ are given by:

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

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

The key determinant of the optimal organization of a single establishment firm is its size.

Proposition 1. *Given the number of layers of management L of the firm,*

a) the number $n^{\ell}_{0,L}$ and the knowledge $z^{\ell}_{0,L}$ of employees at all below-CEO layers $\ell < L$ and the knowledge of the CEO $\bar{z}_{0,L}$ increase with total output $\tilde{q}$, and

b) *the marginal benefit of CEO time* $\varphi_{0,L}$ *and the marginal production cost* $\xi_{0,L}$ *increases with total output* $\tilde{q}$.

c) *The cost function* $C_{0,L}(\tilde{q})$ *strictly increases with total output* $\tilde{q}$. *The average cost function* $AC_{0,L}(\tilde{q})$ *is convex in* $\tilde{q}$. *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.3.1. □

Intuitively, the number of production workers $n_{0,L}^0$ and the CEO knowledge $\bar{z}_{0,L}$ increase because labor and knowledge are complementary inputs in production, so the firm optimally employs a higher amount of both to achieve higher output. An increase in the number of production workers implies that more problems are generated. Thus, more unsolved problems are communicated to higher layers. A higher output therefore leads to an increase in the number of employees $n_{0,L}^\ell$ at all below-CEO layers. The time of the CEO is fixed and does not adjust. Consequently, the knowledge of the employees at the highest below-CEO layer $z_{0,L}^{L-1}$ increases. Otherwise, the CEO could not listen to all problems that are communicated to him. As the lower-layer knowledge levels are recursive functions of higher-layer knowledge, the knowledge at lower layers $z_{0,L}^\ell$, $\ell = 0, ..., L-2$ increases, though to a lesser extent, thereby mitigating the increase in the number of employees at layers $\ell = 1, ..., L-1$.

Larger firms generate more problems, more of which have to be solved at below-CEO layers to meet the CEO's time constraint. The larger the firm, the more beneficial it would therefore be to increase CEO time. That is, the shadow price of the CEO time constraint—the marginal benefit of CEO time—increases with total output. This key implication of the model implies that the limitation to CEO time becomes more and more costly as the firm grows. The marginal production costs increase because higher levels of knowledge at all layers increase the production costs.

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. On the other hand, the quasi-fixed costs of the CEO and the middle managers are spread over a larger output. For quantities below the minimum efficient scale $\tilde{q} < \tilde{q}_L^*$, the latter effect dominates, and for quantities above $\tilde{q} > \tilde{q}_L^*$, the former effect dominates. At the minimum efficient scale, the firm reaches the minimum average costs. The results in Proposition 1 are consistent with the results derived for a knowledge hierarchy with non-overlapping knowledge levels and limited CEO time in Caliendo and Rossi-Hansberg (2012).

The number of managerial layers is determined by equation (3.5). The minimum average cost for a given number of layers decreases and the level of output that achieves this minimum $\tilde{q}_L^*$ increases with the number of layers. The average cost curves of an organization with L and $L+1$ layers cross in the interval $(\tilde{q}_L^*, \tilde{q}_{L+1}^*)$, and the firm adds a layer of management at the crossing (as in Caliendo and Rossi-Hansberg, 2012, Proposition 2). We denote the quantity at the crossing $\tilde{q}_L^{L+1}$. Figure 3.2a illustrates the average cost function of a single establishment firm with only a CEO ($L = 1$) or a CEO and middle managers ($L = 2$).

Adding a layer of middle managers releases CEO time. The middle managers solve part of the problems that are generated by the production workers. They reduce the number of problems sent to the CEO. They thus reduce the marginal benefit of CEO time, i.e. the costs related to the CEO time constraint.

3.4.3 The optimal organization of multi-establishment firms

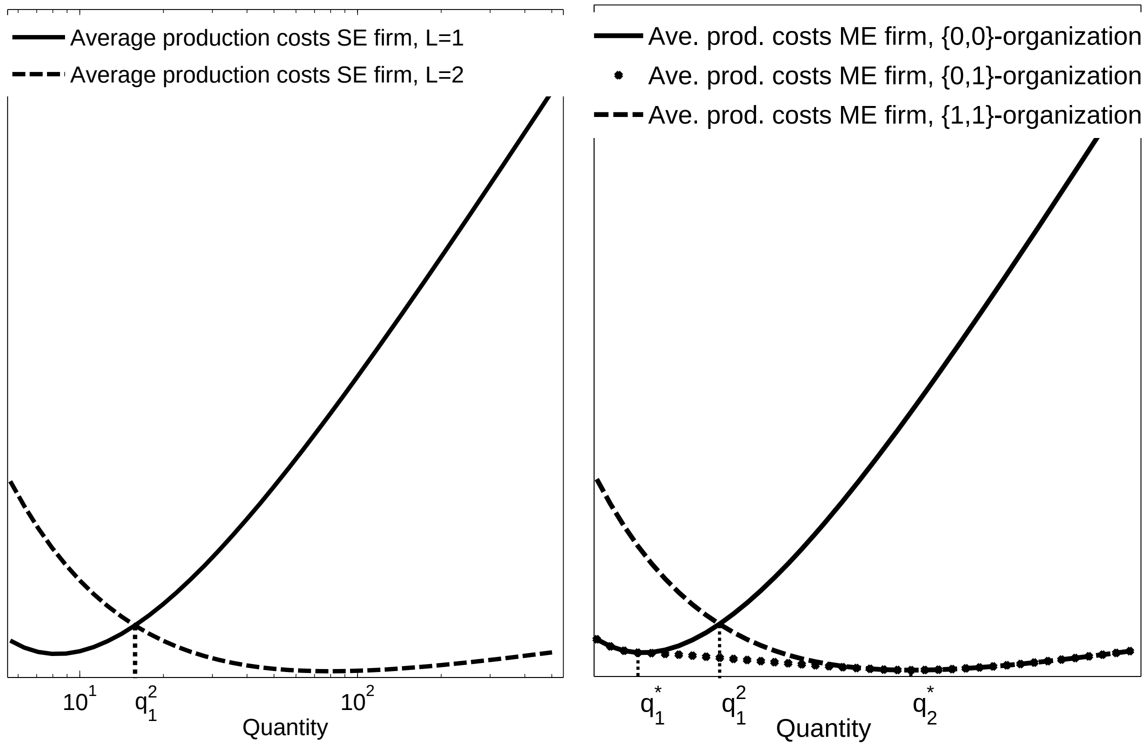
We study the optimal organization of multi-establishment firms in two steps. First, we allow the firm to hire employees in the separate labor markets at both locations, but assume that there is a single output market. As we consider a firm in location 0, the CEO is located in the headquarter establishment in 0. Second, we assume that the firm needs to incur the iceberg-type transport costs $\tau > 1$ to ship output from one location to the other.

Single product market

The firm chooses whether to produce in one establishment at either location or two establishments at both locations as well as the number of managerial layers. We use the term "organizational structure" and the variable ω to denote the number of establishments and number of layers per establishment. All other endogenous variables depend on the location of an establishment and the organizational structure, so we index them by j, ω.

Figure 3.2: Illustration of the average cost functions

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



The figure illustrates the average cost functions of the single and multi-establishment firm for $w_0 = w_1$, $\theta_{00} = \theta_{10}$. Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{10} = \theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = w_1 = 1$. (a): The average cost function of a single establishment firm is U-shaped for a given number of layers $L = 1$, $L = 2$. The firm adds a layer at the intersection $\tilde{q}_1^2$. (b): The average cost function of a multi-establishment firm with a symmetric number of below-CEO layers $\{0, 0\}$ or $\{1, 1\}$ coincides with the average cost function of a single establishment firm. The firm adds a layer at one establishment at the minimum efficient scale $\tilde{q}_1^*$ and a layer at the other establishment at a quantity $\tilde{q} > \tilde{q}_1^2$.

If the firm produces in two establishments, it chooses how much output $q_{j,\omega}$ and which share $s_{j,\omega}$ of CEO time to allocate to each 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 and establishment j.

The optimization problem consists of three parts. First, the firm chooses the optimal organizational structure ω to minimize the total production costs given the total output $\tilde{q}$, analogously to choosing the number of layers in the single establishment case:

$$C(\tilde{q}) = \min_{\omega \in \Omega} \tilde{C}_{0,\omega}(\tilde{q}) \tag{3.14}$$

Second, the firm determines how much output $q_{j,\omega}$ and which share of CEO time $s_{j,\omega}$ to allocate to each establishment, and chooses CEO knowledge $\bar{z}_{0,\omega}$ to minimize the production costs of the chosen organizational structure. The production costs consist of the costs at each establishment j and the remuneration of the CEO time that is not used in production.

$$\tilde{C}_{0,\omega}(\tilde{q}) = \min_{\{q_{j,\omega}, s_{j,\omega}\}_{j=0}^{1}, \bar{z}_{0,\omega} \geq 0} \sum_{j=0}^{1} C_{j,\omega}(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0 (1 + c\bar{z}_{0,\omega}) \tag{3.15}$$

$$\text{s.t.} \quad s_{0,\omega} + s_{1,\omega} \leq 1 \tag{3.16}$$

$$q_{0,\omega} + q_{1,\omega} \geq \tilde{q} \tag{3.17}$$

Constraint (3.16) describes that the CEO has only one unit of time. The production quantities have to sum up at least to the total output $\tilde{q}$, as stated in constraint (3.17).

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

$$C_{j,\omega}\left(q_{j,\omega},s_{j,\omega},\bar{z}_{0,\omega}\right)\begin{cases}\overset{q_{j,\omega}>0}{=} & \min_{\{n^{\ell}_{j,\omega},z^{\ell}_{j,\omega}\}_{\ell=0}^{L_j}\geq 0}\sum_{\ell=0}^{L_j} n^{\ell}_{j,\omega}w_j\left(1+cz^{\ell}_{j,\omega}\right)+s_{j,\omega}w_0\left(1+c\bar{z}_{0,\omega}\right)\\ \overset{q_{j,\omega}=0}{=} & 0\end{cases} \tag{3.18}$$

$$\text{s.t.}\quad n^{0}_{j,\omega}\left(1-e^{-\lambda\bar{z}_{0,\omega}}\right)\geq q_{j,\omega} \tag{3.19}$$

$$s_{j,\omega}\geq n^{0}_{j,\omega}\theta_{j0}e^{-\lambda z^{L_j}_{j,\omega}} \tag{3.20}$$

$$n^{\ell}_{j,\omega}\geq n^{0}_{j,\omega}\theta_{jj}e^{-\lambda z^{\ell-1}_{j,\omega}}\quad\forall\ell=1,...,L_j \tag{3.21}$$

$$\bar{z}_{0,\omega}\geq z^{L_j}_{j,\omega},\quad z^{\ell}_{j,\omega}\geq z^{\ell-1}_{j,\omega}\quad\forall\ell=1,...,L_j \tag{3.22}$$

L_j denotes the number of layers of management at the establishment below the CEO. The constraints (3.19)-(3.22) are analogous to the constraints (3.7)-(3.10).

We solve the problem by backward induction. We first determine the number of employees and their knowledge per layer and establishment, taking as given the firm level choices as well as the organizational structure. We then solve for the knowledge of the CEO, the allocation of his time and of output given the organizational structure, which we determine in the last step. Appendix C.3.2 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 allocation of output and CEO time—through the binding constraints (3.19)-(3.21). The formal expressions are variants of those for the single establishment outcomes in section 3.4.2, which is why we state them in Appendix C.3.2.

Constraint (3.19) determines the number of production workers that depends on the allocated output and CEO knowledge. Constraint (3.20) fixes the knowledge level of the highest layer at the establishment as a function of the allocated share of CEO time and the number of production workers. The knowledge levels of the production workers and managers at lower layers are a recursive function of the knowledge level at the highest layer. Constraint (3.21) determines the number of middle managers as a function of the number of production workers and knowledge levels. 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 an establishment.

Firm-level choices. The firm optimally uses the full unit of CEO time and produces only the required quantity, i.e. the constraints (3.16) and (3.17) are binding. If the firm produces at two establishments, it can reduce the production costs by reallocating total output or CEO time as long as the marginal production costs or the marginal benefit of CEO time are not equal. This key insight drives many of the model's implications.

Proposition 2. *Suppose the firm produces at two establishments. The firm allocates output to equate the marginal production costs across establishments and CEO time to equate the marginal benefit of CEO time across establishments. Formally, in optimum,*

$$\xi_{0,\omega} = \xi_{1,\omega} \quad \textit{and} \tag{3.23}$$

$$\varphi_{0,\omega} = \varphi_{1,\omega}. \tag{3.24}$$

Proof. See Appendix C.3.2. □

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

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

Proof. See Appendix C.3.2. □

Intuitively, the firm only produces at both locations if some advantage at location $j = 1$ counterbalances the higher communication costs across space θ_{10}. The advantage can consist of lower wages or a different managerial structure of the establishment.

Concerning CEO knowledge, the firm balances its marginal benefit and marginal cost, analogously to the single establishment case:

$$w_0 c = \frac{\lambda e^{-\lambda \bar{z}_{0,\omega}}}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \sum_{j=0}^{1} \xi_{j,\omega} q_{j,\omega}. \tag{3.25}$$

Comparative statics with respect to total output $\tilde{q}$. We first determine the impact of total output $\tilde{q}$ on multi-establishment firm organization.

Proposition 3. *Suppose the firm produces at two establishments. Suppose further that there is some asymmetry between the establishments, i.e., either $\theta_{10} > \theta_{00}$, or $w_1 \neq w_0$, or $L_1 \neq L_0$. Given the organizational structure ω,*

a) the total number of employees at all below-CEO layers $\sum_{j=0}^{1} n_{j,\omega}^{\ell}$, $\forall \ell < L$ and CEO knowledge $\bar{z}_{0,\omega}$ increase with total output $\tilde{q}$, while the knowledge of the employees at all below-CEO layers $z_{j,\omega}^{\ell}$ is constant,

b) the share of CEO time $s_{j,\omega}$, and the number of employees at all below-CEO layers $n_{j,\omega}^{\ell}$ at the location with relatively lower (higher) wages increase (decrease) with total output $\tilde{q}$, where the threshold ratio of wages depends on ω. Local output $q_{j,\omega}$ increases if the share of CEO time $s_{j,\omega}$ does, and

c) the marginal benefit of CEO time $\varphi_{j,\omega}$ does not vary and the marginal production cost $\xi_{j,\omega}$ decreases with total output $\tilde{q}$.

d) The cost function $C_{0,\omega}(\tilde{q})$ is strictly increasing with total output $\tilde{q}$.

For full symmetry, i.e., $\theta_{10} = \theta_{00}$, $w_1 = w_0$ and $L_1 = L_0$, the total output has the same effect on the choices of a multi-establishment firm as on those of a single establishment firm.

Proof. See Appendix C.3.2. □

As in a single establishment firm, a higher total output $\tilde{q}$ leads to a higher total number of production workers and a higher CEO knowledge because labor and knowledge are complementary inputs in production. The higher number of production workers leads to a higher number of employees at all below-CEO layers.

The impact of higher output on the other endogenous variables is quite different as long as there is some asymmetry between establishments. The asymmetry may stem from differences in location characteristics, asymmetric numbers of layers or both. Unlike in the single establishment firm, the knowledge levels do not vary with total output. If the firm produces at two establishments, it maintains the same marginal production costs and the same marginal benefit of CEO time across them. For each organizational structure, only one combination of knowledge levels ensures both. Therefore, the below-CEO knowledge levels do not vary with the quantity produced.

The allocation of total output and CEO time reflects that the firm leverages the asymmetries between the establishments. Maintaining two asymmetric establishments effectively allows the firm to produce with two different "production functions", because the firm uses labor and knowledge in different ways in the two establishments. The firm optimally combines the production functions by allocating CEO time and total output. The optimal combination changes with firm size. The larger the firm is, the more important are low wages relative to low communication costs, because the firm hires more employees. The firm thus allocates higher shares of total output and CEO time to the establishment with relatively low wages as it grows. The threshold ratio of wages depends on the organizational structure ω because the impact of lower wages on the production costs depends on the number of layers at the establishments. The number of employees at an establishment depends on the local production quantity and thus varies like it.

As the marginal benefit of CEO time only depends on the below-CEO knowledge levels, it is constant. In contrast, higher CEO knowledge decreases the marginal cost of production, which therefore decreases with total output. The marginal production costs are positive, so the cost function increases with total output. The average production costs are decreasing. This property results because below-CEO knowledge levels are constant. Thus, the costs per unit of labor input are constant. However, CEO knowledge increases with total output. Therefore, more output is producible for every unit of labor input, which leads to the decrease of average costs.

If the two establishments are fully symmetric with respect to both location characteristics and the number of layers, the multi-establishment firm makes the same choices as a single establishment firm. Consequently, changes in total output affect multi-establishment firm organization as stated in Proposition 1 for single establishment firms.

Organizational structure for $w_0 = w_1$, $\theta_{10} = \theta_{00}$. The firm chooses the organizational structure with the minimal production costs. We therefore compare the average production costs of different organizational structures of the multi-establishment firm. L_j denotes the number of below-CEO layers at establishment j and $L = \max_j\{L_j\} + 1$ denotes the number of layers of the firm including the CEO. To simplify the exposition, we first consider the optimal organization when both wages and communication costs are equal, $w_0 = w_1$ and $\theta_{10} = \theta_{00}$.

Proposition 4. *Suppose that wages and communication costs are equal: $w_0 = w_1, \theta_{10} = \theta_{00}$. Let "$\{L_j/L_j\}$-organization" denote the organizational structure of a multi-establishment firm with L_j below-CEO layers at both establishments. Let "$\{L_j/L_j+1\}$-organization" denote the organizational structure of a multi-establishment firm with L_j below-CEO layers at establishment j and $L_j + 1$ below-CEO layers at establishment $k \neq j$.*

a) The average costs of the $\{L_j/L_j\}$-organization coincide with the average costs of a single establishment firm with L layers characterized in Proposition 1c): The average cost function of the $\{L_j/L_j\}$-organization is U-shaped and reaches a minimum at $\tilde{q}^*_L$.

b) The average costs of the $\{L_j/L_j+1\}$-organization are lower than the average costs of the $\{L_j/L_j\}$-organization for output levels $\tilde{q} > \tilde{q}^*_L$.

c) The average cost function of the $\{L_j+1/L_j+1\}$-organization intersects the average cost function of the $\{L_j/L_j\}$-organization at the output $\tilde{q}^{L+1}_L$, with $q^*_{L+1} > \tilde{q}^{L+1}_L > q^*_L$. It intersects the average cost function of the $\{L_j/L_j+1\}$-organization at the output $\tilde{q} > \tilde{q}^{L+1}_L$.

As a result, the multi-establishment firm with L_j below-CEO layers at both establishments adds a layer of management at one establishment at the output $\tilde{q}^*_L$ and a layer at the other establishment at a output $\tilde{q} \in (\tilde{q}^{L+1}_L, \tilde{q}^*_{L+1})$.

Proof. See Appendix C.3.2. □

Figure 3.2b 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 costs of the {0/0}-organization increase for quantities above the minimum efficient scale q_1^*, whereas the average costs of the {0/1}-organization decrease, as stated in parts a) and b) of Proposition 4. Consequently, the former intersect the average costs of the {1/1}-organization at a lower quantity than the latter, illustrating part c).[11]

Proposition 4 is a key result of the model. It states that the multi-establishment firm reorganizes gradually. A single establishment firm with L_j below-CEO layers adds a managerial layer at the size $\tilde{q}_L^{L+1}$. The multi-establishment firm with L_j below-CEO layers adds a layer at one establishment at output $\tilde{q}_L^* < \tilde{q}_L^{L+1}$. It adds a layer at the other establishment at output $\tilde{q} > \tilde{q}_L^{L+1}$.

This difference results because the multi-establishment organization is free to allocate total output and CEO time. The quantity q_L^* is the minimum efficient scale of the $\{L_j/L_j\}$-organization, because the organization has the minimum average costs at that quantity. A multi-establishment firm with a $\{L_j/L_j+1\}$-organization would allocate total output to the establishment with L_j below-CEO layers at q_L^*. For quantities $\tilde{q} > q_L^*$, the average costs of the $\{L_j/L_j\}$-organization increase. The average costs of the multi-establishment firm with the $\{L_j/L_j+1\}$-organization decrease, because it can allocate a share of total output to the establishment with L_j+1 below-CEO layers. For quantities close to the minimum efficient scale, only a small share is allocated to the establishment with L_j+1 below-CEO layers, but the larger the quantity $\tilde{q}$, the larger its share of production. The additional managerial layer releases CEO time that is allocated to the establishment with L_j below-CEO layers. The additional managerial layer thus increases efficiency at both establishments. In consequence, the multi-establishment firm only switches to the $\{L_j+1/L_j+1\}$-organization at output $\tilde{q} > \tilde{q}_L^{L+1}$.

The flexible allocation of output and CEO time explains why the multi-establishment firm does not increase the knowledge at below-CEO layers when it grows, as explained in Proposition 3a). The marginal benefit of CEO time is constant and the marginal cost

[11] The average costs of the {0/1}-organization coincide with the average costs of the the {0/0}-organization and the {1/1}-organization for quantities below and above their minimum efficient scales, because in these ranges, single establishment production is more efficient than production with the {0/1}-organization.

of production decreases at both establishments in the multi-establishment firm.

Proposition 4 contains an important insight about the optimal organization of the multi-establishment firm: The organization of its establishments is interdependent. The number of layers at one establishment depends on the number of layers at the other establishment.

Organizational structure for $w_0 > w_1$, $\theta_{10} > \theta_{00}$. The communication costs across space θ_{10} only affect the multi-establishment firm organization.

Proposition 5. *Suppose the firm produces at two establishments. Suppose further that $\theta_{10} > \theta_{00}$ and that w_1 is low relative to w_0. Given the organizational structure ω,*

a) the total number of employees at all below-CEO layers $\sum_{j=0}^{1} n_{j,\omega}^{\ell}$, $\forall \ell < L$ decreases with the communication costs θ_{10}, while the knowledge of the CEO $\bar{z}_{0,\omega}$ and the knowledge of the employees at all below-CEO layers $z_{j,\omega}^{\ell}$, $\forall \ell < L$ increase, and

b) the marginal benefit of CEO time $\varphi_{j,\omega}$ and the marginal production cost $\xi_{j,\omega}$ increase with the communication costs θ_{10}.

The increase of the below-CEO knowledge levels with the communication costs θ_{10} is stronger at higher than at lower layers and at the subordinate establishment than at the headquarters.

Proof. See Appendix C.3.2. □

An increase in the communication costs θ_{10} implies that it is more costly to use the CEO's knowledge. Generating problems is more costly, because they may have to be communicated to the CEO to produce output. The firm therefore hires fewer production workers and thus fewer middle managers at all below-CEO layers. To nevertheless produce the required amount of output, the firm adjusts the optimal levels of knowledge. The firm increases CEO knowledge to maintain total output despite the lower number of production possibilities due to the lower number of workers. The firm also increases the knowledge at the below-CEO levels because of the CEO time constraint: With higher communication costs, the CEO spends more time listening to a given communication, so more problems have to be solved by the below-CEO layers. These adjustments lead to higher marginal costs and a higher marginal benefit of CEO time. The increase of knowledge is stronger at higher layers because the number of employees is lower at higher layers, so it is cheaper to increase their knowledge.

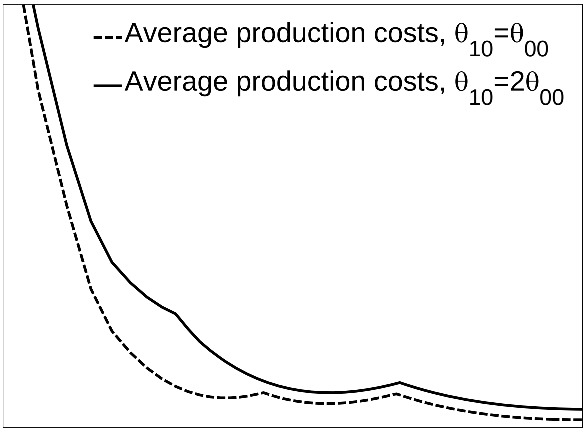
Proposition 5 again illustrates the interdependence of multi-establishment firm organization: Changes of parameters that affect one establishment lead to organizational adjustments at both establishments because the firm maintains equal marginal costs and marginal benefit of CEO time across establishments. Higher communication costs θ_{10} imply that accessing CEO knowledge is more expensive for employees in location $j = 1$. Their knowledge increases, so they communicate fewer problems. As the problem probability distribution function is downward sloping, the marginal product of knowledge is decreasing. The firm optimally compensates for the additional communication costs by not only increasing the knowledge at the subordinate establishment, but also at the headquarters. The requirement to equate both marginal production costs and marginal benefit of CEO time (Proposition 2) implies that the increase is higher at the subordinate establishment than at the headquarters.

Higher communication costs θ_{10} increase the average production costs. They thus modify, but do not fundamentally alter the insights about the optimal organizational structure in Proposition 4. The higher θ_{10} is, the higher the average production costs are, and the smaller the range of quantities is for which multi-establishment production is efficient. Still, the multi-establishment firm reorganizes gradually.

Separate product markets

Assuming that the firm freely allocates both total output and CEO time across establishments makes it easier to analytically solve the model. It implies that the below-CEO knowledge levels do not vary with total output (Proposition 3). However, firms are likely limited in their flexibility to allocate total output in practice. We therefore extend the model to separate local output markets. We assume that the firm has to pay an iceberg-type transport cost τ if it sells output produced in location j in location $k \neq j$. The firm faces two possibly different levels of local demand, $\{\tilde{q}_j\}_{j=0}^{1}$. The additional assumption does not affect the optimization problem at the establishment level outlined in equations (3.18)-(3.22), nor does it change the choice of the optimal organizational

Figure 3.3: Illustration of the average cost functions, local demand



The figure illustrates the average cost function of a multi-establishment firm for $w_0 = w_1$, $\theta_{00} \leq \theta_{10}$. At each kink, the multi-establishment firm adds a layer at one establishment. The communication costs affect the quantity at which the firm reorganizes.

structure in equation (3.14). However, it alters the optimal allocation of output and CEO time:

$$\tilde{C}_{0,\omega}\left(\{\tilde{q}_j\}_{j=0}^{1}\right) = \min_{\{q_{j,\omega},s_{j,\omega}\}_{j=0}^{1},\bar{z}_{0,\omega}\geq 0} \sum_{j=0}^{1} C_{j,\omega}\left(q_{j,\omega}, s_{j,\omega}, \bar{z}_{0,\omega}\right) + \left[1 - \sum_{j=0}^{1} s_{j,\omega}\right] w_0\left(1 + c\bar{z}_{0,\omega}\right) \tag{3.26}$$

$$\text{s.t.} \quad s_{0,\omega} + s_{1,\omega} \leq 1 \tag{3.27}$$

$$\mathbb{1}(q_{j,\omega} \geq \tilde{q}_j \wedge q_{k,\omega} \leq \tilde{q}_k)(q_{j,\omega} - \tilde{q}_j + \tau(q_{k,\omega} - \tilde{q}_k)) \geq 0, \quad k \neq j \tag{3.28}$$

Constraint (3.28) states that output at a location j $q_{j,\omega}$ has to be large enough to satisfy demand at j and a possible difference between local production and local demand at location k including transport cost: $q_{j,\omega} \geq \tilde{q}_j + \tau(q_{k,\omega} - \tilde{q}_k)$.

The transport friction leads to different marginal production costs across locations.

Proposition 6. *Suppose the firm produces at two establishments and incurs transport costs $\tau > 1$ to ship output from one location to the other. The firm allocates CEO time to equate the marginal benefit of CEO time across establishments:*

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

The firm allocates output to equate the marginal production costs, adjusted by the transport costs, across establishments if possible. Formally, in optimum,

$$\xi_{0,\omega} = \tau^{-1}\xi_{1,\omega} \qquad \textit{if } q_{0,\omega} = \tilde{q}_0 + \tau(q_{1,\omega} - \tilde{q}_1), \tag{3.29}$$

$$\xi_{0,\omega} = \tau\xi_{1,\omega} \qquad \textit{if } q_{1,\omega} = \tilde{q}_1 + \tau(q_{0,\omega} - \tilde{q}_0), \ \textit{and} \tag{3.30}$$

$$\tau^{-1}\xi_{1,\omega} < \xi_{0,\omega} < \tau\xi_{1,\omega} \qquad \textit{if } q_{1,\omega} = \tilde{q}_1 \wedge q_{0,\omega} = \tilde{q}_0. \tag{3.31}$$

Proof. See Appendix C.3.2. □

Proposition 6 distinguishes three cases. The firm may produce output at location 0 to satisfy local demand and part of the demand at location 1 (equation 3.29). In this case, it allocates output such that the marginal production costs at location 1 are equal to the marginal production costs at location 0 adjusted by the transport costs. Reversely, the firm equates the marginal production costs at location 0 and the marginal production costs at location 1 adjusted by the transport costs if it produces more output than demanded at 1 (equation 3.30). However, this allocation is only feasible under certain parameter conditions. The firm produces output strictly for local demand if the marginal costs at one location cannot be equalized to the transport adjusted marginal costs at the other (equation 3.31). If the marginal production costs at one location adjusted for transport costs are lower than the marginal costs at the other location, the firm produces total output at this location, i.e. is a single establishment firm.

The allocation of output across locations has implications for the comparative statics results of the model. If the firm equates the marginal production costs adjusted for transport costs (i.e., if either (3.29) or (3.30) hold), the results from Proposition 3 apply. In particular, the below-CEO knowledge levels do not vary with output. Otherwise, the comparative statics results are similar to those for a single establishment firm in Proposition 1. In particular, the below-CEO knowledge levels increase with higher output and the average cost function is U-shaped. Intuitively, if the firm cannot freely allocate output, it faces similar constraints as a single establishment firm: The firm allocates the CEO time to equate its marginal benefit across establishments. It then chooses the number and knowledge level of employees in order to minimize the production costs of local demand, taking as given the allocated share of CEO time. Figure 3.3 illustrates the resulting average cost function.

Figure 3.3 also illustrates that the communication costs across space θ_{10} affect the number of managerial layers of the firm in a setting with local demand. The higher the communication costs are, the smaller is the quantity at which the firm adds a layer at one establishment, as a comparison of the solid and dashed lines show. Higher communication costs increase the knowledge levels of employees and thus the marginal production costs. Adding a layer helps the firm to mitigate the cost increase, because it allows decreasing production worker knowledge and thus marginal costs. In contrast, the firm avoids the high communication costs across space by producing total output at one establishment if it can freely allocate output (see Corollary 1).

Importantly, the main implications of the model in section 3.4.3 are unchanged. In both settings, the organization of establishments is interdependent: As the firm allocates CEO time to equate its marginal benefit across establishments, changes to the communication costs θ_{10} and other parameters affect the number and knowledge levels of employees at both establishments. Likewise, the number of layers at an establishment depends on the managerial structure of the other establishment. In both settings, the multi-establishment firm reorganizes gradually. The firm adds a layer at one establishment at a smaller size than if it were a single establishment firm and at the other establishment at a larger size.

3.4.4 The optimal total output

For simplicity, we endogenize the optimal total output in the setting with a single output market studied in section 3.4.3.

We consider a setting with many firms i that each produce a differentiated product. Agents maximize their utility (3.4) subject to their budget constraint. The total demand results from multiplying the individual demand by the number of agents per location:

$$q(\alpha_i) = \alpha_i(R_0 + R_1)P^{\sigma-1}p(x(\alpha_i))^{-\sigma}$$

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

Each firm chooses the optimal output to maximize profits given the taste draw:

$$\max_{\tilde{q}\geq 0} \pi_i(\alpha_i) = p(\tilde{q}(\alpha_i))\tilde{q}(\alpha_i) - C(\tilde{q}) \tag{3.32}$$

Substituting the demand function and solving for the optimal output yields

$$\tilde{q}(\alpha_i) = \alpha_i \left(R_0 + R_1\right) \left(\frac{\sigma}{\sigma - 1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma}, \tag{3.33}$$

where we make explicit that the marginal costs $\xi_{0,\omega}$ are a function of α_i through output. The optimal price is a constant mark-up over marginal costs:

$$p(\alpha_i) = \frac{\sigma}{\sigma - 1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right) \tag{3.34}$$

Proposition 7. *The optimal output $\tilde{q}(\alpha_i)$ increases continuously with the taste parameter α_i. If the firm produces at both locations, the optimal output decreases with the communication costs across space θ_{10}.*

Proof. See Appendix C.3.3. □

This result implies that firm geography affects firm size. The higher the communication costs across space are, the higher the marginal production costs are and in consequence, the lower total output is. Higher communication costs thus decrease total profits and the incentive to produce at both locations.

3.4.5 Summary

As stated at the beginning of this section, the objective of the model is to explain why the number of establishments of a firm affects its managerial organization. Taking stock, we find that the model is consistent with the three facts uncovered in section 3.3.

Proposition 4 shows that multi-establishment firms add a managerial layer at one establishment at a lower output and at the other establishment at a larger output than single establishment firms. This is consistent with Facts 2 and 3. Multi-establishment firms are predicted to have more managerial layers than single establishment firms and to reorganize gradually.

Extending the model with transport frictions helps understand why the number of layers of a multi-establishment firm increases with establishment distance, as found in Fact 2. Proposition 6 implies that the marginal costs differ across establishments depending on the transport costs. The extended model is therefore consistent with recent evidence from the literature that distance to the headquarters decreases establishment performance (e.g., Giroud, 2013). It explains the negative effect of distance on the location probability and establishment size uncovered in Fact 1.

The key driver of the model implications is that firms allocate the common resource CEO time such that its marginal benefit is equal across establishments. Consequently, the organizational structure is interdependent across establishments. We next use an exogenous change of spatial frictions to provide evidence for this model prediction.

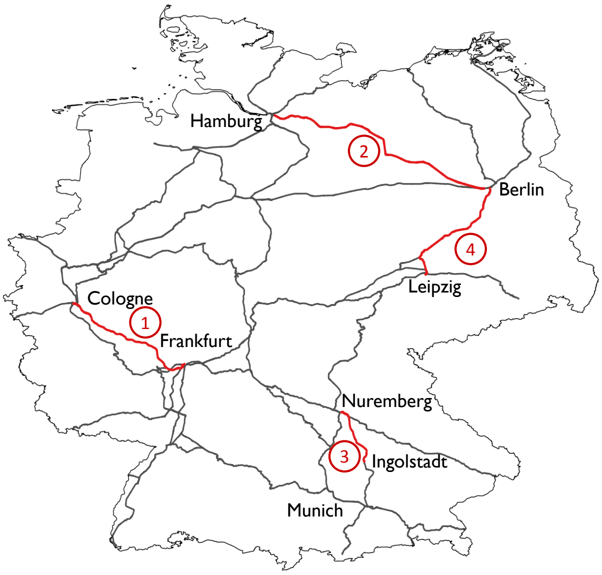
3.5 Organizational response to new high speed train routes

Section 3.5 traces the organizational response of multi-establishment firms to an exogenous change of the economic environment. We exploit the opening of new high-speed train routes between major cities in Germany (similar to Charnoz et al., 2015; Bernard et al., 2017). The high speed train routes make it easier for managers to travel between the headquarters and subordinate establishments and thus facilitate face-to-face interactions within multi-establishment firms. They thus provide a plausibly exogenous reduction of the costs to manage subordinate establishments from the headquarters. In the vocabulary of the model, they decrease θ_{10}, the communication costs across space.

3.5.1 Travel time changes due to new high speed train routes

We use information on the changes in the travel times between German cities due to the introduction of four high-speed train routes. Deutsche Bahn AG, the state-owned German railway firm, either constructed new rails (routes 1, 3, 4) or substantially upgraded the existing railway network (route 2). Route 1 almost halved the travel time between Frankfurt and Cologne from 135 minutes to 76 minutes. Service started in August 2002

Figure 3.4: The new high speed train routes and the high speed railway network



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

(Eurailpress.de, 2002). Route 2 reduced the travel time between Hamburg and Berlin from 135 minutes to 90 minutes from December 2004 (Eurailpress.de, 2004). Deutsche Bahn AG launched train service on Route 3, the new train route between Ingolstadt and Nuremberg in May 2006. The route reduced the travel time between the two cities from 66 minutes to 30 minutes (Brux, 2006). Route 4 decreased the travel time between Leipzig and Berlin from 145 minutes to 75 minutes in the same year (Eurailpress.de, 2006). Figure 3.4 shows a map of the new high speed train routes and how they connect to the existing high speed rail network.

Trains on all routes exclusively transport people. 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 high speed rail network. Between the connected cities, it is faster to take the train than the car or even plane (if one accounts for waiting times at the airport). In fact, regular plane service between Cologne Bonn Airport and Stuttgart Airport was discontinued in 2002,[12] and the service between Cologne Bonn Airport and Frankfurt Airport was discontinued in 2007.[13] The number

[12]It takes about 80 minutes to travel from Frankfurt to Stuttgart by train.

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

of flights between Cologne Bonn Airport and Nuremberg Airport dropped substantially (Deutscher Bundestag, 2007).

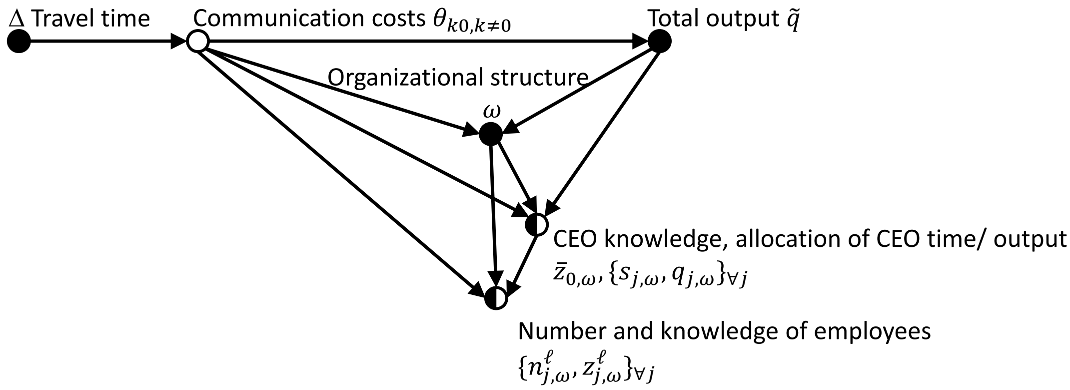
As figure 3.4 shows, the German railway network is very complex compared to other countries. Paris, for instance, is the center of the French railway network that has approximately a "star" network structure. In contrast, there are many different connections between medium sized cities in the German railway network. The travel time reductions therefore propagate through the network and affect more locations than only those at the immediate ends. For example, Route 3 between Ingolstadt and Nuremberg decreased travel times from Munich to many medium sized cities such as Würzburg or Bamberg, Leipzig, and, together with Route 4, Berlin.

To capture the impact of the train routes, we obtain information on the mean and minimum travel times in the years 2000, 2004 and 2008 from the Deutsche Bahn AG. Our data comprise 115 train stations that are connected to the ICE network, the German high speed train network, in at least one of the three years. To make sure that temporary construction works do not affect the travel times, Deutsche Bahn AG computed the mean travel times based on information from three different weekdays in March, June and November. Travel times may change over time for several reasons, including adjustments of the time table, construction works, new changeover connections, or new high speed routes. To allow us to disentangle lower travel times due to the new routes and other reasons, the data contain an indicator for station pairs where more than 50% of passengers used one of the new high speed routes in 2008. We merge the travel times and the data on multi-establishment firms using the information on the county where the establishment is located.

3.5.2 Model predictions

Section 3.4 shows that the parameter θ_{10} is an important determinant of multi-establishment firm organization. It affects the organization of the establishments and CEO knowledge (Proposition 5) and has an impact on the organizational structure (section 3.4.3). Importantly, it affects the production quantity (Proposition 7), another key determinant of multi-establishment firm organization (Proposition 3, 4). We take this complexity into account in our empirical analysis.

Figure 3.5: Response of endogenous variables to change in the travel times



The graph illustrates the response of the endogenous variables to a change of the travel time according to the model in Section 3.4. The arrows denote causal relationships between the variables at the nodes. The node symbol • (∘) denotes that a variable is (un)observable. ◐ denotes that part of a group of variables is observable and the other part is unobservable. The change of the travel time affects the communication costs between a subordinate establishment k and the headquarters $\theta_{k0,k\neq 0}$. θ_{k0} has a direct effect on the managerial organization and affects the organizational structure ω, CEO knowledge $\bar{z}_{0,\omega}$, the allocation of CEO time and output $\{s_{j,\omega}, q_{j,\omega}\}_{\forall j}$, and the number and knowledge of employees $n^{\ell}_{j,\omega}, z^{\ell}_{j,\omega}$. $\theta_{k0,k\neq 0}$ has an indirect effect on these variables through the impact on total output $\tilde{q}$.

Figure 3.5 illustrates the model implications using a directed graph. Solid circles denote observable variables and hollow circles denote unobservable variables. The arrows denote causal links between the variables at the nodes. To keep the graph simple, we group variables and use semi-solid circles if only part of the group is observable.

The figure shows that changes to the communication costs between an establishment k and the headquarters $\theta_{k0,k\neq 0}$ have direct effects on the organization of multi-establishment firms and indirect effects. First, changes to θ_{k0} affect the optimal total output $\tilde{q}$. θ_{k0} has a direct effect on the organizational structure ω and an indirect effect through $\tilde{q}$ because total output is a determinant of ω. The firm level choices CEO knowledge $\bar{z}_{0,\omega}$, the allocation of CEO time $s_{j,\omega}$ and the allocation of output $q_{j,\omega}$ depend directly on θ_{k0}, but also indirectly through $\tilde{q}$ and ω. Similarly, the establishment level choices concerning the number and knowledge of employees per layer $n^{\ell}_{j,\omega}, z^{\ell}_{j,\omega}$ depend directly on θ_{k0} and indirectly through $\bar{z}_{0,\omega}, s_{j,\omega}, q_{j,\omega}$ and ω.

The figure makes clear that the changes to the travel times provide an exogenous change of the model parameter θ_{k0} that affects many endogenous variables. Put differently, we have more endogenous variables than exogenous instruments. In result, we can only estimate the total, direct plus indirect, effect of θ_{k0} on the organization of firms.

We cannot test the predictions of Proposition 5, for example, because the Proposition derives the impact of θ_{k0} on endogenous outcomes taking as given total output and the organizational structure. Only Proposition 7 is testable, because output only directly depends on θ_{k0}. Nevertheless, our empirical exercise is informative about the main model prediction: The key insight of the model is that changes to the economic environment of one establishment affect the organizational structure of all establishments of a multi-establishment firm. We take two predictions to the data:

Prediction 1. *Firm size increases after a reduction in the travel time between a subordinate establishment and the headquarters of the multi-establishment firm.*

Prediction 2. *A reduction in the travel time between one subordinate establishment and the headquarters of the multi-establishment firm affects the organization of all establishments.*

To provide evidence for these predictions, we estimate three specifications. At the firm level, we estimate:

$$y_{it} = \delta_0 + \delta_1 D_{\exists j \text{ s.t. } \theta\downarrow, it} + \alpha_i + \alpha_{ct} + \epsilon_{it} \tag{3.35}$$

i refers to a multi-establishment firm, j to a subordinate establishment, c to the headquarters county and t indexes time. y_{ijt} denotes the outcome variables. The main variable of interest is the indicator variable $D_{\exists j \text{ s.t. } \theta\downarrow, it}$. It is equal to one if the travel time between at least one subordinate establishment and the headquarters decreases. α_i is a firm fixed effect. α_{ct} is a county-year fixed effect.

With view to Figure 3.5, we choose the following outcome variables: We use the number of employees to capture total output and the allocation of output across establishments. We employ the number of hierarchical layers to reflect the organizational structure ω of the firm. The share of employees in managerial occupations captures both the number of hierarchical layers and the allocation of employees and their knowledge across layers. Given Prediction 1, we expect $\delta_1 > 0$ for firm size.

The specification mimics a differences-in-differences estimation equation. The "treatment" is faster travel time between at least one subordinate establishment and the headquarters. Its baseline effect is captured by the firm fixed effect. The time fixed effect captures the "after" dummy. The indicator variable $D_{\exists j \text{ s.t. } \theta\downarrow,it}$ is equivalent to the interaction term of the "treatment" and "after" dummy variables.

To understand the impact of the "treatment" on the establishment outcomes, we estimate:

$$y_{ijt} = \beta_0 + \beta_1 D_{\theta\downarrow,ijt} + \alpha_j + \alpha_{ct} + \epsilon_{ijt} \tag{3.36}$$

where the main variable of interest is $D_{\theta\downarrow,ijt}$, an indicator variable for lower travel times between the establishment and its headquarters. α_j is an establishment fixed effect. We use the same outcome variables as for the firm level.

As the model predicts that changes of the communication costs between one subordinate establishment and the headquarters may affect the organization of the other establishments, we exclude "non-treated" subordinate establishments of firms with at least one treated subordinate establishment from the control group. The regressions estimate how outcomes of subordinate establishments with a lower travel time to the headquarters evolve compared to subordinate establishments that belong to firms where none of the subordinate establishments is treated. In the baseline regressions, we also restrict the sample to establishments at locations connected to the high speed rail network to avoid that unobservable differences between establishments connected and not connected to the network drive the results.

We adjust the estimation equation to explore the impact on non-treated subordinate establishments of treated firms following Prediction 2:

$$y_{ikt} = \gamma_0 + \gamma_1 D_{\exists j \neq k \text{ s.t. } \theta\downarrow,ikt} + \alpha_k + \alpha_{ct} + \epsilon_{ikt}, \quad k \neq j \tag{3.37}$$

k refers to an untreated subordinate establishment. The indicator variable $D_{\exists j \neq k \text{ s.t. } \theta\downarrow,ikt}$ is equal to one if the travel time between one of the other establishments of the firm and the headquarters decreased by at least 30 minutes. The control group consists of subordinate establishments of non-treated firms.

In our baseline specification, we set the indicators $D_{\theta\downarrow,ijt}$, $D_{\exists j \neq k \text{ s.t. } \theta\downarrow,ikt}$ and $D_{\exists j \text{ s.t. } \theta\downarrow,it}$ equal to one if the travel time between the subordinate establishment j and the headquarters decreases by at least 30 minutes. The high-speed train routes decrease the travel times by at least 30 minutes. As Appendix Table C.11 shows, virtually none of

Table 3.8: Regression results, firm size

	Firm	Treated establishment			Untreated est.	
Number of employees	(1)	(2)	(3)	(4)	(5)	(6)
$D_{\exists j \text{ s.t. } \theta\downarrow,it}$	0.045**					
	(0.013)					
$D_{\theta\downarrow,ijt}$		−0.008	−0.000	0.077*		
		(0.030)	(0.059)	(0.032)		
$D_{\exists j\neq k \text{ s.t. } \theta\downarrow,ikt}$					−0.209***	−0.213***
					(0.030)	(0.030)
R-squared	0.976	0.920	0.920	0.874	0.859	0.860
Firm FE	Y	N	N	N	N	N
Establishment FE	N	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y
Observations	11,218	12,210	4,086	24,489	20,314	22,416

2000-2010 panel, only firms with at least 10 employees. Standard errors clustered at county level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* number of employees of a firm (column 1) or subordinate establishment (columns 2-6). *Independent variables:* $D_{\exists j \text{ s.t. } \theta\downarrow,it}$: 1 if travel time to headquarters is reduced by at least 30 minutes at one establishment of firm. $D_{\theta\downarrow,ijt}$: 1 if travel time to headquarters is reduced by at least 30 minutes. $D_{\exists j\neq k \text{ s.t. } \theta\downarrow,ikt}$ 1 if travel time to headquarters is reduced by at least 30 minutes at another establishment of firm. Column 3: subordinate establishments with maximum travel time to headquarters of 150 minutes in 2000.

the non-high-speed-route connections exhibit a decrease in travel times of 30 minutes or more. The threshold thus helps us to ensure that the reduction is indeed driven by the exogenous new route instead of potentially endogenous adjustments to the time table that may respond to changes of demand. We exclude subordinate establishments with minor travel time reductions from the control group. We conduct robustness checks to show that our insights are robust to these restrictions.

One may be worried that the difference in the travel times also affects other model parameters, such as local wages because employees commute longer distances. The empirical methodology isolates the impact of lower face-to-face frictions on firm organization from the effect of other economic forces. Lower local wages benefit establishments of treated and untreated firms, so our estimation strategy differences out their effect. In addition, the county-year fixed effects capture the local economic conditions.

3.5.3 Regression results

We present our regression results by outcome variable. Table 3.8 presents the regression results for size. The first column refers to the results for the firm. Columns 2-4 refer to the treated establishment. Column 2 contains the baseline specification. Column 3 restricts the sample to establishments with initial travel time of at most 2.5 hours, because we expect the train routes to reduce management costs to a larger degree at shorter distances, where round-trips are feasible within a day. Column 4 includes non-treated establishments of the same firm in the control group. Columns 5 and 6 present the results for the untreated establishments. In column 5, we restrict the control group to establishments of firms with at least three establishments to make it more comparable to the treated group. In column 6, we drop this restriction.

Consistent with Prediction 1, we find that treated firms grow compared to untreated firms. The reduction of travel time thus seems to enhance firm efficiency. The growth rate of treated establishments is not significantly different from the growth rate of establishments of untreated firms (columns 2, 3). However, treated establishments grow significantly faster than the control group if we include untreated establishments of the firm. This indicates that employment is reallocated within the firm towards the treated establishment. The negative impact of faster train routes on the untreated establishments in columns 5 and 6 is consistent with this interpretation.

Table 3.9 contains the regression results for the organizational structure. At the firm level, we find that the managerial share tends to increase, but the effect is not significantly different from zero (column 1). The number of layers increases (column 2). This finding indicates that the positive impact of a larger size on the number of layers derived in Proposition 3 outweighs the negative direct effect of the communication costs.

At the establishment level, the managerial share of treated establishments decreases compared to the managerial share of subordinate establishments of untreated firms. The estimates in column 3 are equivalent to a reduction by 10% at the mean. The decrease is considerably stronger when we restrict the sample to subordinate establishments with a maximum travel time to the headquarters of 2.5 hours in 2000 in column 4. In contrast, including the untreated subordinate establishments of a treated firm in column 5 mitigates the effect. It is still negative, but only marginally significant (P-value 13.0%). The

Table 3.9: Regression results, managerial share

Managerial share/ # layers	Firm (1)	Firm (2)	Treated establishment (3)	Treated establishment (4)	Treated establishment (5)	Untreated est. (6)	Untreated est. (7)
$D_{\exists j \text{ s.t. } \theta\downarrow,it}$	0.266	0.045^{*}					
	(0.217)	(0.019)					
$D_{\theta\downarrow,ijt}$			-1.572^{*}	-2.832^{**}	-0.917		
			(0.706)	(0.842)	(0.606)		
$D_{\exists j\neq k \text{ s.t. } \theta\downarrow,ikt}$						-1.336^{**}	-1.398^{**}
						(0.471)	(0.463)
R-squared	0.964	0.887	0.846	0.846	0.813	0.803	0.804
Firm FE	Y	Y	N	N	N	N	N
Establishment FE	N	N	Y	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y	Y	Y	Y
Observations	11,218	11,218	12,210	4,086	24,489	20,314	22,416

2000-2010 panel, only firms with at least 10 employees. Standard errors clustered at county level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* share of employees in managerial occupations of firm (column 1) or subordinate establishment (columns 3-7); # layers of firm (column 2). *Independent variables:* see Table 3.8. Column 3: subordinate establishments with maximum travel time to headquarters of 150 minutes in 2000.

managerial share at the untreated subordinate establishments decreases for both definitions of the control group. In unreported regressions, we find that the share of managers at the headquarters increases, in particular at levels 2 and 3 (coefficient all management layers .307 (s.e. .242), levels 2 and 3 $.543^{*}$ (s.e. .240)). This finding reconciles the difference between the establishment level management shares and the headquarters.

The main take-away from the table is that changes in the economic environment of one establishment affect the managerial share of all establishments of the firm, consistent with Prediction 2. Several factors may contribute to the estimated negative effect. Lower communication costs decrease the need for local management, as illustrated in section 3.4.3. Further, larger firm size leads to higher knowledge levels at all layers according to Proposition 3, which additionally decreases the ratio of managers to workers.

Robustness. We conduct a number of robustness checks. First, we restrict the sample to the years 2000, 2004, and 2008 for which we have travel time information. Second, we include establishments with travel time reduction of less than 30 minutes in the control group. Third, we add establishments in counties without an ICE train station. Finally, we redefine treatment as a travel time reduction of at least 10 minutes. Appendix

Table C.12 shows that our results are robust to these modifications for the treated establishments. Throughout, the effect of lower travel time on the managerial share is negative and significant (if only marginally so in column 3, P-value of 11.3%).

3.6 Conclusion

This chapter shows that geographic frictions between one subordinate establishment and the headquarters affect the managerial organization of all establishments of a multi-establishment firm. Our insights imply that the impact of local economic conditions on multi-establishment firms goes beyond their impact on specific establishments. Multi-establishment firms account for disproportionate shares of aggregate output and employment. Understanding how local conditions propagate through the network of multi-establishment firms is an exciting and relevant area for future research.

Appendix A

Appendix to Chapter 1

A.1 Descriptive statistics and illustrations

A.1.1 Descriptive statistics

Table A.1: Descriptive statistics of non-manufacturing firms with donut, 2000

	0.5-2 km			2.8-4.2 km			difference	
	Mean	SD	N	Mean	SD	N	(4) - (1)	SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Performance								
Log(sales, survey)	14.11	1.88	536	13.88	1.72	120	−0.22	0.19
Log(value added, survey)	13.37	1.94	467	13.23	1.80	107	−0.14	0.20
Share of purchased inputs	0.43	0.23	474	0.43	0.20	111	0.01	0.02
Log(sales per FTE)	12.41	1.60	513	12.32	1.50	115	−0.09	0.16
Log(sales per empl.)	11.17	0.73	509	11.11	0.73	115	−0.06	0.08
Employment size								
Log(full-time empl., SIAB)	1.66	1.49	536	1.59	1.45	120	−0.07	0.15
Log(total empl., survey)	3.16	1.67	536	3.00	1.57	120	−0.16	0.17
Log(daily wage sum, SIAB)	5.87	1.66	536	5.81	1.60	120	−0.06	0.17
Log(monthly wage sum, survey)	10.34	1.99	516	10.20	1.84	117	−0.15	0.20
Skill composition								
Share low-skilled	0.07	0.18	536	0.05	0.15	120	−0.01	0.02
Share medium-skilled	0.81	0.29	536	0.87	0.22	120	0.06	0.03
Share high-skilled	0.10	0.23	536	0.08	0.16	120	−0.03	0.02

This table shows the descriptive statistics of the firms in panel sample 2000 to 2011 in the first year. Data sources are reported in the table. The shares of skill groups are calculated based on the composition in SIAB.

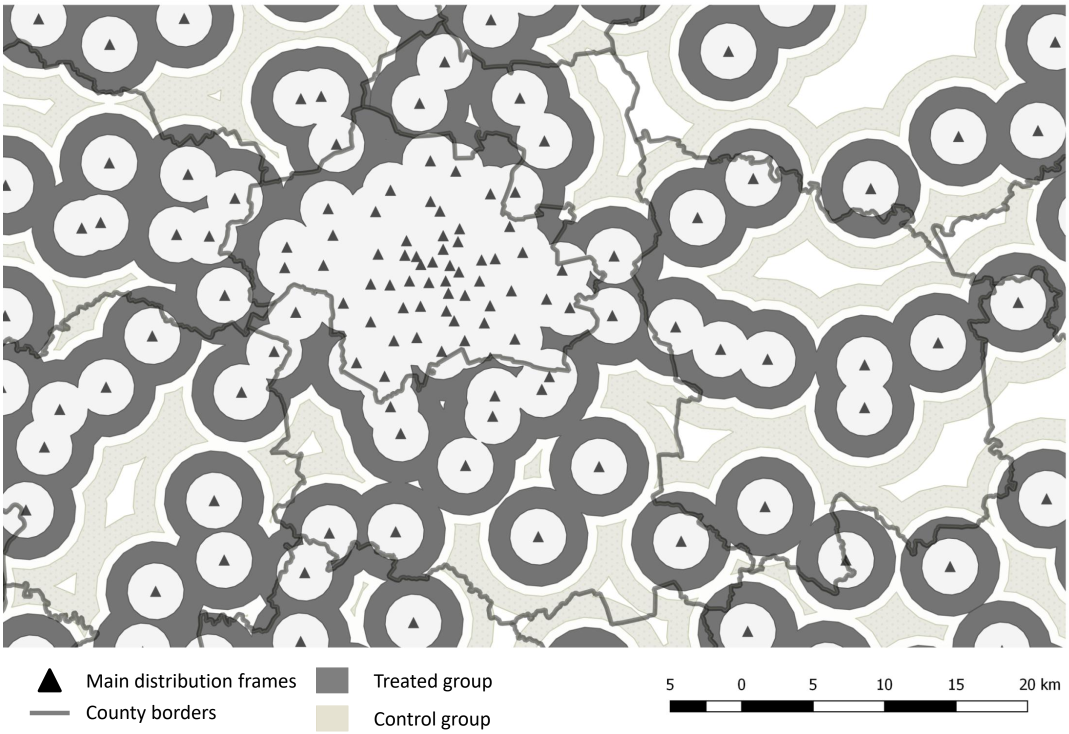
Table A.2: Descriptive statistics of non-manufacturing firms without donut, 2000

	0.5-2 km			2-3.5 km			difference	
	Mean	SD	N	Mean	SD	N	(4) - (1)	SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Performance								
Log(Sales, survey)	14.13	1.91	577	14.08	1.76	233	−0.05	0.15
Log(Value added, survey)	13.39	1.95	501	13.45	1.84	207	0.06	0.16
Share of purchased inputs	0.43	0.23	509	0.42	0.21	212	−0.01	0.02
Log(Sales per FTE)	12.43	1.63	552	12.32	1.48	224	−0.10	0.13
Log(Sales per empl.)	11.17	0.76	548	11.11	0.73	223	−0.05	0.06
Employment size								
Log(Full-Time Empl., SIAB)	1.66	1.50	577	1.75	1.53	233	0.08	0.05
Log(Total Empl., survey)	3.19	1.71	577	3.12	1.53	233	−0.06	0.13
Log(Daily wage sum, SIAB)	5.87	1.67	577	5.95	1.73	233	0.07	0.13
Log(Monthly wage sum, survey)	10.37	2.03	556	10.37	1.82	228	−0.00	0.15
Skill composition								
Share Low Skilled	0.07	0.18	577	0.05	0.16	233	−0.02	0.01
Share Medium Skilled	0.81	0.29	577	0.84	0.25	233	0.03	0.02
Share High Skilled	0.11	0.24	577	0.10	0.21	233	−0.01	0.02

This table shows the descriptive statistics of the firms in panel sample 2000 to 2011 in the first year. Data sources are reported in the table. The shares of skill groups are calculated based on the composition in SIAB.

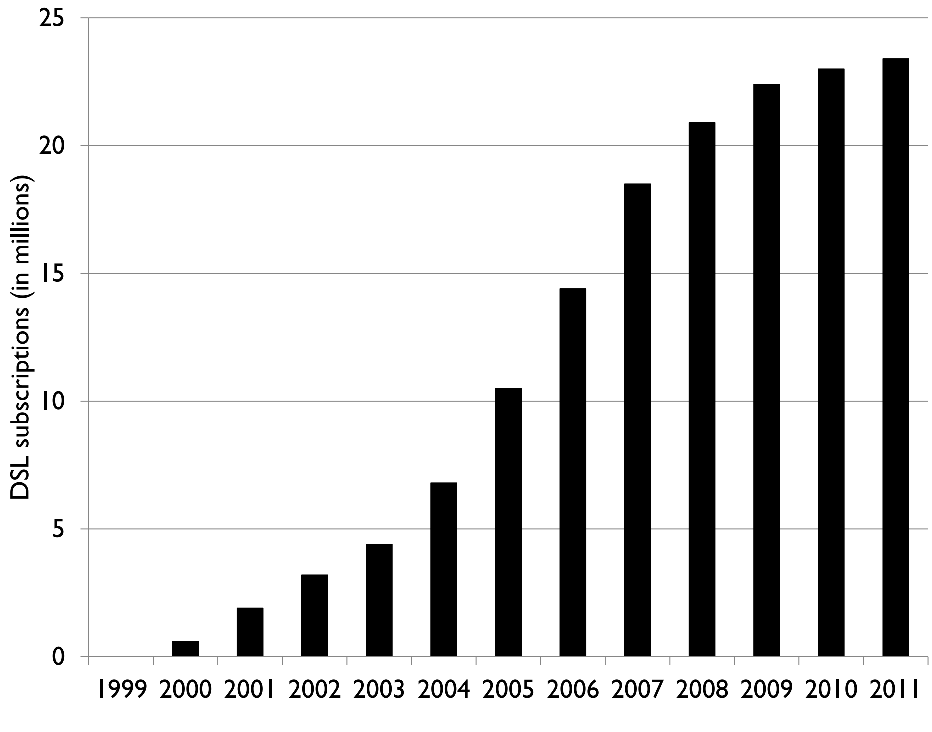
A.1.2 Identification strategy

Figure A.1: Map of main distribution frames



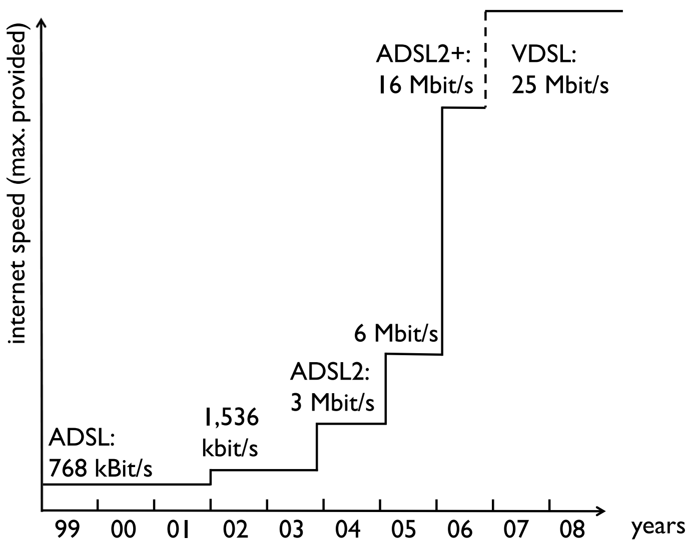
The figure illustrates the definition of treatment and control group for the analysis of speed upgrades. Grey lines show county borders. Black triangles resemble MDFs. Grey circles resemble the areas of treated firms. Light circles resemble the areas of firms in the control group.

Figure A.2: Number of DSL subscriptions in Germany



The figure shows the number of DSL subscriptions in Germany over time (Bundesnetzagentur, 2011)

Figure A.3: Timeline



The figure shows the timeline of the introduction and upgrading of broadband internet technology following Schnabel (2015). In 2000, customers received ADSL with a maximum download speed of 768 kBit/s. In later years, the speed was upgraded up to 6 MBit/s. In 2006, however, ADSL2+ provided a much larger speed upgrades with up to 16 MBit/s provided by most telecoms companies. Further, VDSL was introduced but required large infrastructure investments which took several years. Own illustration.

A.2 Results

A.2.1 Results of introduction of the broadband internet

Table A.3: Robustness checks on timing of introduction

Dependent variable		D_{post} from 2000	Excl. early municip.
Performance			
Log(revenues)		−0.018	−0.012
Source: survey		(0.077)	(0.081)
	N	1,255	1,201
Log(value added)		−0.015	0.033
Source: survey		(0.108)	(0.107)
	N	987	937
Log(revenues per full-time		0.086	0.132
employee (FTE))		(0.111)	(0.113)
Sources: survey and SIAB	N	1,123	1,073
Log(revenues per employee)		0.144**	0.147**
Source: survey		(0.063)	(0.061)
	N	1,121	1,071
Employment			
Log(total employment)		−0.084	−0.066
Source: survey		(0.068)	(0.073)
	N	1,255	1,201
Log(full-time employment)		−0.077	−0.094
Source: SIAB		(0.108)	(0.114)
	N	1,255	1,201
Log(wage sum)		−0.118	−0.128
Source: SIAB		(0.117)	(0.122)
	N	1,255	1,201
Log(wage sum)		−0.147*	−0.152*
Source: survey		(0.077)	(0.084)
	N	1,216	1,162
Skill composition			
Share of low-skilled		−0.020	−0.024
Source: SIAB		(0.015)	(0.019)
	N	1,255	1,201
Share of medium-skilled		0.033*	0.046**
Source: SIAB		(0.019)	(0.022)
	N	1,255	1,201
Share of high-skilled		−0.014	−0.024
Source: SIAB		(0.016)	(0.017)
	N	1,255	1,201

This table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment period and 2001 to 2005 as the treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_1 \,\mathsf{DSL\ access}_i\, D_{post,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{post,t}$ is an indicator variable for the treatment period from 2001 on. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results for β_1. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The regression in the first column defines the post period from 2000 on. The sample in the second column excludes municipalities in which any MDF was upgraded to the first generation of broadband internet in 2000.

Table A.4: Yearly effects of the introduction of broadband internet: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{1997}$	−0.033	−0.050	−0.019	−0.107
	(0.142)	(0.086)	(0.149)	(0.109)
$D_{DSL} \times D_{1998}$	0.007	−0.013	0.020	−0.049
	(0.127)	(0.083)	(0.134)	(0.100)
$D_{DSL} \times D_{1999}$	0.021	−0.009	−0.007	−0.066
	(0.123)	(0.086)	(0.126)	(0.105)
$D_{DSL} \times D_{2000}$	−0.017	−0.094	−0.053	−0.134
	(0.123)	(0.083)	(0.126)	(0.100)
$D_{DSL} \times D_{2001}$	0.009	−0.103	−0.034	−0.230**
	(0.127)	(0.086)	(0.132)	(0.106)
$D_{DSL} \times D_{2002}$	0.011	−0.005	−0.015	−0.084
	(0.141)	(0.093)	(0.143)	(0.112)
$D_{DSL} \times D_{2003}$	−0.153	−0.100	−0.194	−0.305***
	(0.142)	(0.088)	(0.146)	(0.110)
$D_{DSL} \times D_{2004}$	−0.219	−0.175*	−0.272*	−0.332***
	(0.154)	(0.104)	(0.156)	(0.121)
$D_{DSL} \times D_{2005}$	−0.224	−0.197	−0.286	−0.198
	(0.176)	(0.154)	(0.183)	(0.152)
R-squared	0.937	0.965	0.948	0.967
Obs.	1,255	1,255	1,255	1,216

This table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment and 2001 to 2005 as treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\ \mathsf{DSL\ access}_i\ D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data. The sample only contains non-manufacturing firms.

Table A.5: Yearly effects of the introduction of broadband internet: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{1997}$	0.029	0.053	0.005	0.031	0.174
	(0.084)	(0.143)	(0.046)	(0.159)	(0.122)
$D_{DSL} \times D_{1998}$	0.040	0.060	−0.009	0.058	0.140
	(0.079)	(0.147)	(0.043)	(0.143)	(0.123)
$D_{DSL} \times D_{1999}$	−0.010	−0.171	0.078*	−0.008	0.207
	(0.078)	(0.141)	(0.043)	(0.142)	(0.126)
$D_{DSL} \times D_{2000}$	0.015	−0.067	0.039	0.012	0.200
	(0.083)	(0.136)	(0.040)	(0.144)	(0.129)
$D_{DSL} \times D_{2001}$	0.001	−0.050	0.017	0.088	0.275**
	(0.088)	(0.167)	(0.044)	(0.146)	(0.130)
$D_{DSL} \times D_{2002}$	0.047	−0.058	0.031	0.083	0.354***
	(0.088)	(0.143)	(0.043)	(0.159)	(0.135)
$D_{DSL} \times D_{2003}$	−0.030	−0.003	−0.020	0.147	0.260*
	(0.093)	(0.149)	(0.046)	(0.159)	(0.134)
$D_{DSL} \times D_{2004}$	−0.098	−0.097	0.005	0.138	0.278*
	(0.107)	(0.170)	(0.049)	(0.168)	(0.148)
$D_{DSL} \times D_{2005}$	0.019	0.116	−0.005	0.288+	0.496***
	(0.128)	(0.177)	(0.049)	(0.167)	(0.157)
R-squared	0.973	0.937	0.662	0.934	0.839
Obs.	1,255	987	1,002	1,123	1,121

This table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment and 2001 to 2005 as treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\ \mathsf{DSL\ access}_i\ D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. The sample only contains non-manufacturing firms.

Table A.6: Main regression results of the introduction of broadband internet

Dependent variable		Pooled	Non-manuf.	Manuf.
Performance				
Log(revenues)		−0.006	−0.023	0.004
Source: survey		(0.031)	(0.040)	(0.047)
	N	2,184	1,255	929
Log(value added)		0.029	0.007	0.034
Source: survey		(0.054)	(0.068)	(0.087)
	N	1,734	987	747
Log(revenues per full-time		0.096**	0.116*	0.070
employee (FTE))		(0.044)	(0.060)	(0.064)
Sources: survey and SIAB	N	1,971	1,123	848
Log(revenues per employee)		0.118***	0.140***	0.088**
Source: survey		(0.036)	(0.052)	(0.049)
	N	1,967	1,121	846
Employment				
Log(total employment)		−0.074*	−0.091	−0.068
Source: survey		(0.041)	(0.056)	(0.061)
	N	2,184	1,255	929
Log(full-time employment)		−0.075***	−0.071*	−0.100***
Source: SIAB		(0.027)	(0.038)	(0.037)
	N	2,184	1,255	929
Log(wage sum)		−0.101**	−0.125**	−0.086
Source: SIAB		(0.043)	(0.057)	(0.064)
	N	2,184	1,255	929
Log(wage sum)		−0.079**	−0.150***	−0.013
Source: survey		(0.037)	(0.047)	(0.060)
	N	2,114	1,216	898
Skill composition				
Share of low-skilled		−0.015**	−0.020**	−0.006
Source: SIAB		(0.006)	(0.009)	(0.007)
	N	2,184	1,255	929
Share of medium-skilled		0.017	0.043***	−0.020
Source: SIAB		(0.012)	(0.013)	(0.022)
	N	2,184	1,255	929
Share of high-skilled		−0.001	−0.022**	0.025
Source: SIAB		(0.011)	(0.010)	(0.021)
	N	2,184	1,255	929

This table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment period and 2001 to 2005 as the treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_1 \, \mathsf{DSL\ access}_i \, D_{post,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{post,t}$ is an indicator variable for the treatment period from 2001 on. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results for β_1. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.2.2 Main results of speed upgrades

Table A.7: Main results on speed upgrades: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.008	0.120	−0.004	0.019	−0.043
	(0.055)	(0.088)	(0.024)	(0.067)	(0.055)
$D_{DSL} \times D_{2002}$	0.028	0.124	−0.016	−0.018	−0.035
	(0.052)	(0.087)	(0.023)	(0.073)	(0.059)
$D_{DSL} \times D_{2003}$	0.128**	0.210**	−0.009	0.120	0.061
	(0.053)	(0.100)	(0.025)	(0.079)	(0.060)
$D_{DSL} \times D_{2004}$	0.168***	0.296***	−0.025	0.108	0.083
	(0.057)	(0.096)	(0.026)	(0.074)	(0.060)
$D_{DSL} \times D_{2005}$	0.138**	0.222**	−0.035	0.063	0.039
	(0.060)	(0.097)	(0.025)	(0.081)	(0.067)
$D_{DSL} \times D_{2006}$	0.147**	0.304***	−0.048*	0.072	0.037
	(0.061)	(0.099)	(0.029)	(0.082)	(0.069)
$D_{DSL} \times D_{2007}$	0.191***	0.324***	−0.021	0.069	0.058
	(0.064)	(0.106)	(0.031)	(0.078)	(0.069)
$D_{DSL} \times D_{2008}$	0.150**	0.192*	0.005	0.075	0.034
	(0.061)	(0.106)	(0.029)	(0.077)	(0.068)
$D_{DSL} \times D_{2009}$	0.158**	0.242**	−0.009	0.076	0.105
	(0.063)	(0.103)	(0.032)	(0.080)	(0.067)
$D_{DSL} \times D_{2010}$	0.161**	0.255**	−0.008	0.072	0.093
	(0.063)	(0.103)	(0.032)	(0.081)	(0.064)
$D_{DSL} \times D_{2011}$	0.168**	0.277***	−0.039	0.084	0.112
	(0.066)	(0.101)	(0.029)	(0.093)	(0.072)
R-squared	0.973	0.944	0.699	0.925	0.822
Obs.	6,274	5,054	5,189	5,472	5,455

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data.

Table A.8: Main results on speed upgrades: firm size and skill composition

	log(employment)		log(wage sum)		skill shares		
	full-time	total	daily	survey	low	med.	high
$D_{DSL} \times D_{2001}$	−0.019	0.069	0.015	0.070	−0.015	0.001	0.013
	(0.055)	(0.049)	(0.056)	(0.066)	(0.010)	(0.018)	(0.016)
$D_{DSL} \times D_{2002}$	0.046	0.065	0.052	0.066	−0.007	0.013	−0.006
	(0.058)	(0.049)	(0.057)	(0.066)	(0.009)	(0.019)	(0.017)
$D_{DSL} \times D_{2003}$	0.017	0.078	0.047	0.094	−0.005	0.003	0.005
	(0.061)	(0.052)	(0.060)	(0.069)	(0.009)	(0.021)	(0.018)
$D_{DSL} \times D_{2004}$	0.062	0.119**	0.101	0.205***	−0.008	0.026	−0.018
	(0.064)	(0.058)	(0.064)	(0.074)	(0.011)	(0.019)	(0.016)
$D_{DSL} \times D_{2005}$	0.059	0.132**	0.120*	0.191***	0.005	0.007	−0.008
	(0.062)	(0.058)	(0.064)	(0.074)	(0.010)	(0.020)	(0.018)
$D_{DSL} \times D_{2006}$	0.083	0.118**	0.107*	0.123*	−0.008	0.030*	−0.018
	(0.061)	(0.056)	(0.062)	(0.072)	(0.009)	(0.017)	(0.015)
$D_{DSL} \times D_{2007}$	0.087	0.100*	0.110*	0.143*	−0.006	0.042**	−0.031*
	(0.061)	(0.055)	(0.061)	(0.075)	(0.010)	(0.021)	(0.018)
$D_{DSL} \times D_{2008}$	0.069	0.124**	0.090	0.198**	−0.001	0.040*	−0.039*
	(0.067)	(0.058)	(0.067)	(0.079)	(0.010)	(0.023)	(0.020)
$D_{DSL} \times D_{2009}$	0.066	0.107*	0.114*	0.198**	−0.016	0.054**	−0.037*
	(0.065)	(0.059)	(0.066)	(0.083)	(0.011)	(0.022)	(0.020)
$D_{DSL} \times D_{2010}$	0.077	0.103*	0.083	0.118	−0.025*	0.059***	−0.031*
	(0.070)	(0.059)	(0.073)	(0.077)	(0.015)	(0.023)	(0.018)
$D_{DSL} \times D_{2011}$	0.078	0.131**	0.115	0.102	−0.019	0.067**	−0.045*
	(0.091)	(0.066)	(0.091)	(0.077)	(0.015)	(0.030)	(0.026)
R-squared	0.951	0.978	0.960	0.971	0.762	0.809	0.833
Obs.	6,274	6,274	6,274	6,068	6,274	6,274	6,274

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum retrieved from survey data. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree. The sample only contains non-manufacturing firms.

A.2.3 Results of speed upgrades in manufacturing

Table A.9: Results on speed upgrades, manufacturing: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.008	0.120	−0.004	0.019	−0.043
	(0.055)	(0.088)	(0.024)	(0.067)	(0.055)
$D_{DSL} \times D_{2002}$	0.028	0.124	−0.016	−0.018	−0.035
	(0.052)	(0.087)	(0.023)	(0.073)	(0.059)
$D_{DSL} \times D_{2003}$	0.128**	0.210**	−0.009	0.120	0.061
	(0.053)	(0.100)	(0.025)	(0.079)	(0.060)
$D_{DSL} \times D_{2004}$	0.168***	0.296***	−0.025	0.108	0.083
	(0.057)	(0.096)	(0.026)	(0.074)	(0.060)
$D_{DSL} \times D_{2005}$	0.138**	0.222**	−0.035	0.063	0.039
	(0.060)	(0.097)	(0.025)	(0.081)	(0.067)
$D_{DSL} \times D_{2006}$	0.147**	0.304***	−0.048*	0.072	0.037
	(0.061)	(0.099)	(0.029)	(0.082)	(0.069)
$D_{DSL} \times D_{2007}$	0.191***	0.324***	−0.021	0.069	0.058
	(0.064)	(0.106)	(0.031)	(0.078)	(0.069)
$D_{DSL} \times D_{2008}$	0.150**	0.192*	0.005	0.075	0.034
	(0.061)	(0.106)	(0.029)	(0.077)	(0.068)
$D_{DSL} \times D_{2009}$	0.158**	0.242**	−0.009	0.076	0.105
	(0.063)	(0.103)	(0.032)	(0.080)	(0.067)
$D_{DSL} \times D_{2010}$	0.161**	0.255**	−0.008	0.072	0.093
	(0.063)	(0.103)	(0.032)	(0.081)	(0.064)
$D_{DSL} \times D_{2011}$	0.168**	0.277***	−0.039	0.084	0.112
	(0.066)	(0.101)	(0.029)	(0.093)	(0.072)
R-squared	0.973	0.944	0.699	0.925	0.822
Obs.	6,274	5,054	5,189	5,472	5,455

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\,D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data.

Table A.10: Results on speed upgrades, manufacturing: firm size

	log(employment)			log(wage sum)	
	full-time	total	prod.	daily	survey
$D_{DSL} \times D_{2001}$	−0.019	0.069	−0.043	0.015	0.070
	(0.055)	(0.049)	(0.059)	(0.056)	(0.066)
$D_{DSL} \times D_{2002}$	0.046	0.065	0.015	0.052	0.066
	(0.058)	(0.049)	(0.062)	(0.057)	(0.066)
$D_{DSL} \times D_{2003}$	0.017	0.078	−0.017	0.047	0.094
	(0.061)	(0.052)	(0.066)	(0.060)	(0.069)
$D_{DSL} \times D_{2004}$	0.062	0.119**	0.036	0.101	0.205***
	(0.064)	(0.058)	(0.068)	(0.064)	(0.074)
$D_{DSL} \times D_{2005}$	0.059	0.132**	0.049	0.120*	0.191***
	(0.062)	(0.058)	(0.067)	(0.064)	(0.074)
$D_{DSL} \times D_{2006}$	0.083	0.118**	0.050	0.107*	0.123*
	(0.061)	(0.056)	(0.066)	(0.062)	(0.072)
$D_{DSL} \times D_{2007}$	0.087	0.100*	0.064	0.110*	0.143*
	(0.061)	(0.055)	(0.065)	(0.061)	(0.075)
$D_{DSL} \times D_{2008}$	0.069	0.124**	0.047	0.090	0.198**
	(0.067)	(0.058)	(0.072)	(0.067)	(0.079)
$D_{DSL} \times D_{2009}$	0.066	0.107*	0.029	0.114*	0.198**
	(0.065)	(0.059)	(0.071)	(0.066)	(0.083)
$D_{DSL} \times D_{2010}$	0.077	0.103*	0.047	0.083	0.118
	(0.070)	(0.059)	(0.081)	(0.073)	(0.077)
$D_{DSL} \times D_{2011}$	0.078	0.131**	−0.120	0.115	0.102
	(0.091)	(0.066)	(0.128)	(0.091)	(0.077)
R-squared	0.951	0.978	0.939	0.960	0.971
Obs.	6,274	6,274	6,274	6,274	6,068

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data.

Table A.11: Results on speed upgrades, manufacturing: shares of skill groups

	low-skill	medium-skill	high-skill
$D_{DSL} \times D_{2001}$	0.004	−0.007	0.004
	(0.008)	(0.015)	(0.012)
$D_{DSL} \times D_{2002}$	−0.000	−0.010	0.010
	(0.008)	(0.014)	(0.011)
$D_{DSL} \times D_{2003}$	−0.000	−0.021	0.020*
	(0.009)	(0.014)	(0.011)
$D_{DSL} \times D_{2004}$	0.012	−0.050***	0.038***
	(0.008)	(0.015)	(0.013)
$D_{DSL} \times D_{2005}$	0.004	−0.049***	0.039***
	(0.008)	(0.015)	(0.013)
$D_{DSL} \times D_{2006}$	0.001	−0.042***	0.037***
	(0.008)	(0.015)	(0.013)
$D_{DSL} \times D_{2007}$	−0.006	−0.034**	0.035***
	(0.008)	(0.015)	(0.012)
$D_{DSL} \times D_{2008}$	−0.003	−0.033**	0.027**
	(0.009)	(0.015)	(0.011)
$D_{DSL} \times D_{2009}$	−0.020*	−0.028*	0.040***
	(0.011)	(0.016)	(0.013)
$D_{DSL} \times D_{2010}$	−0.008	−0.023	0.020
	(0.009)	(0.016)	(0.013)
$D_{DSL} \times D_{2011}$	−0.021	−0.003	0.015
	(0.014)	(0.020)	(0.018)
R-squared	0.824	0.801	0.771
Observations	6,117	6,117	6,117

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\,D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree.

A.2.4 Robustness checks on introduction of the broadband internet

Table A.12: Other robustness checks for the introduction of broadband internet

Dependent variable		(1)	(2)	(3)	(4)	(5)
Performance						
Log(revenues)		−0.006	−0.208	−0.440**	−0.043	0.066
Source: survey		(0.123)	(0.152)	(0.185)	(0.081)	(0.099)
	N	572	381	211	1,179	793
Log(value added)		0.050	−0.073	−0.389*	−0.011	0.172
Source: survey		(0.168)	(0.233)	(0.194)	(0.109)	(0.136)
	N	428	319	164	1,179	626
Log(revenues per full-time		0.161	0.060	−0.028	−0.102	0.207
employee (FTE))		(0.170)	(0.211)	(0.224)	(0.111)	(0.143)
Sources: survey and SIAB	N	506	354	199	1,054	704
Log(revenues per employee)		0.154	0.134	−0.126	0.144**	0.131
Source: survey		(0.110)	(0.087)	(0.255)	(0.061)	(0.083)
	N	505	354	198	1,052	702
Employment						
Log(total employment)		−0.067	−0.217	−0.393	−0.098	−0.036
Source: survey		(0.159)	(0.157)	(0.331)	(0.107)	(0.150)
	N	572	381	211	1,179	793
Log(full-time employment)		−0.061	−0.139	−0.126	−0.087	0.043
Source: SIAB		(0.113)	(0.100)	(0.211)	(0.071)	(0.096)
	N	572	381	211	1,179	793
Log(wage sum)		−0.144	−0.252*	−0.423	−0.127	−0.075
Source: SIAB		(0.190)	(0.151)	(0.316)	(0.112)	(0.158)
	N	572	381	211	1,179	793
Log(wage sum)		−0.174	−0.207*	−0.275	−0.158*	−0.074
Source: survey		(0.129)	(0.110)	(0.296)	(0.082)	(0.104)
	N	553	373	211	1,140	793
Skill composition						
Share of low-skilled		−0.020	0.002	0.026	−0.016	−0.023
Source: SIAB		(0.028)	(0.009)	(0.028)	(0.016)	(0.021)
	N	572	381	211	1,179	793
Share of medium-skilled		0.038	0.043	−0.024	0.033*	0.052*
Source: SIAB		(0.030)	(0.048)	(0.040)	(0.018)	(0.030)
	N	572	381	211	1,179	793
Share of high-skilled		−0.016	−0.051	−0.021	−0.016	−0.029
Source: SIAB		(0.012)	(0.046)	(0.014)	(0.012)	(0.025)
	N	572	381	211	1,179	793

The table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment period and 2001 to 2005 as the treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_1 \, \mathsf{DSL\ access}_i \, D_{post,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{post,t}$ is an indicator variable for the treatment period from 2001 on. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 4.2 and 5.7 km from the MDF. The table reports the results for β_1. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. All samples only contain non-manufacturing firms. Samples are further restricted by robustness check. Column (1) contains firms that did not invest in ICT in 2000. Column (2) contains firms that invested in ICT in 2000. Column (3) contains firms in Western Germany. Column (4) contains firms in municipalities with at least one household with DSL access in 2005. Column (5) contains firms founded in 1992 or before.

A.2.5 Robustness checks using the distance

Table A.13: Results on speed upgrades, distance: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$Distance \times D_{2001}$	−0.016	−0.058	−0.025	−0.049
	(0.040)	(0.037)	(0.040)	(0.048)
$Distance \times D_{2002}$	−0.058	−0.083**	−0.056	−0.065
	(0.042)	(0.037)	(0.041)	(0.049)
$Distance \times D_{2003}$	−0.063	−0.114***	−0.073*	−0.098*
	(0.045)	(0.040)	(0.044)	(0.052)
$Distance \times D_{2004}$	−0.148***	−0.168***	−0.148***	−0.200***
	(0.047)	(0.043)	(0.045)	(0.055)
$Distance \times D_{2005}$	−0.146***	−0.165***	−0.158***	−0.185***
	(0.044)	(0.042)	(0.044)	(0.052)
$Distance \times D_{2006}$	−0.120***	−0.154***	−0.133***	−0.153***
	(0.046)	(0.043)	(0.046)	(0.054)
$Distance \times D_{2007}$	−0.139***	−0.158***	−0.161***	−0.184***
	(0.045)	(0.043)	(0.044)	(0.056)
$Distance \times D_{2008}$	−0.103**	−0.162***	−0.110**	−0.180***
	(0.049)	(0.044)	(0.049)	(0.058)
$Distance \times D_{2009}$	−0.084*	−0.150***	−0.119**	−0.177***
	(0.048)	(0.045)	(0.047)	(0.061)
$Distance \times D_{2010}$	−0.109**	−0.146***	−0.129**	−0.117**
	(0.051)	(0.043)	(0.052)	(0.057)
$Distance \times D_{2011}$	−0.108	−0.144***	−0.124	−0.103*
	(0.076)	(0.046)	(0.077)	(0.060)
Constant	1.851***	2.999***	6.054***	10.181***
	(0.018)	(0.014)	(0.018)	(0.017)
R-squared	0.950	0.977	0.959	0.971
Obs.	6,352	6,352	6,352	6,145

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{log(distance)}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{log(distance)}_i$ is a continuous variable indicating the distance between the firm and the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data.

Table A.14: Results on speed upgrades, distance: performance

	log(revenues)	log(VA)	Shr. of Inputs	log(revenues p. FTE)	log(revenues p. empl.)
$Distance \times D_{2001}$	0.018	−0.046	0.007	0.026	0.060
	(0.045)	(0.067)	(0.017)	(0.051)	(0.044)
$Distance \times D_{2002}$	−0.043	−0.103	0.022	−0.012	0.042
	(0.041)	(0.071)	(0.018)	(0.055)	(0.047)
$Distance \times D_{2003}$	−0.105**	−0.204***	0.016	−0.048	−0.008
	(0.043)	(0.077)	(0.018)	(0.058)	(0.049)
$Distance \times D_{2004}$	−0.116**	−0.226***	0.021	0.008	0.010
	(0.046)	(0.076)	(0.019)	(0.058)	(0.049)
$Distance \times D_{2005}$	−0.126***	−0.220***	0.034*	0.006	0.028
	(0.047)	(0.079)	(0.020)	(0.061)	(0.054)
$Distance \times D_{2006}$	−0.128***	−0.256***	0.054***	−0.012	0.075
	(0.048)	(0.076)	(0.019)	(0.062)	(0.053)
$Distance \times D_{2007}$	−0.124**	−0.252***	0.035*	0.044	0.075
	(0.049)	(0.083)	(0.020)	(0.057)	(0.051)
$Distance \times D_{2008}$	−0.118**	−0.194**	0.016	−0.027	0.057
	(0.048)	(0.082)	(0.020)	(0.059)	(0.053)
$Distance \times D_{2009}$	−0.121**	−0.245***	0.034	−0.033	−0.009
	(0.049)	(0.079)	(0.022)	(0.061)	(0.053)
$Distance \times D_{2010}$	−0.124**	−0.279***	0.044*	−0.035	−0.006
	(0.049)	(0.081)	(0.024)	(0.063)	(0.050)
$Distance \times D_{2011}$	−0.139***	−0.258***	0.055**	−0.003	0.022
	(0.051)	(0.082)	(0.022)	(0.087)	(0.055)
Constant	13.906***	13.235***	0.420***	11.983***	11.160***
	(0.016)	(0.027)	(0.007)	(0.021)	(0.018)
R-squared	0.972	0.943	0.700	0.924	0.813
Obs.	6,352	5,111	5,247	5,538	5,520

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{log(distance)}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{log(distance)}_i$ is a continuous variable indicating the distance between the firm and the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data.

Table A.15: Results on speed upgrades, distance: performance

	low-skill	medium-skill	high-skill
$Distance \times D_{2001}$	0.003	−0.004	0.003
	(0.008)	(0.012)	(0.009)
$Distance \times D_{2002}$	−0.007	−0.010	0.017*
	(0.008)	(0.013)	(0.010)
$Distance \times D_{2003}$	−0.007	−0.003	0.008
	(0.009)	(0.013)	(0.010)
$Distance \times D_{2004}$	−0.010	−0.007	0.018*
	(0.009)	(0.013)	(0.010)
$Distance \times D_{2005}$	−0.019**	−0.000	0.019*
	(0.009)	(0.014)	(0.011)
$Distance \times D_{2006}$	−0.011	−0.007	0.019*
	(0.008)	(0.013)	(0.010)
$Distance \times D_{2007}$	−0.012	−0.000	0.013
	(0.010)	(0.015)	(0.010)
$Distance \times D_{2008}$	−0.025**	0.010	0.022**
	(0.010)	(0.015)	(0.011)
$Distance \times D_{2009}$	−0.007	−0.006	0.014
	(0.010)	(0.015)	(0.012)
$Distance \times D_{2010}$	−0.006	−0.003	0.010
	(0.015)	(0.018)	(0.012)
$Distance \times D_{2011}$	−0.002	−0.021	0.028
	(0.013)	(0.022)	(0.017)
Constant	0.073***	0.815***	0.101***
	(0.004)	(0.006)	(0.004)
R-squared	0.758	0.810	0.835
Observations	6,352	6,352	6,352

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{log(distance)}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{log(distance)}_i$ is a continuous variable indicating the distance between the firm and the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree.

A.2.6 Robustness check on introduction: 2-3.5 km to 3.5-5 km

Table A.16: Regression results comparing 2-3.5 km to 3.5-5 km

	Pooled	Non-manuf.	Manuf.
Performance			
log(revenues)	-0.064^{**}	-0.072^{*}	-0.057
	(0.030)	(0.041)	(0.042)
	2,359	1,381	978
log(value added)	-0.034	-0.107^{*}	0.048
	(0.053)	(0.065)	(0.086)
	1,874	987	792
log(revenues per FTE)	-0.023	-0.020	-0.024
	(0.039)	(0.057)	(0.053)
	2,115	1,123	894
log(revenues per employee)	-0.006	0.016	-0.031
	(0.034)	(0.051)	(0.044)
	2,110	1,121	892
Employment			
log(total employment)	-0.040	-0.038	-0.050
	(0.033)	(0.048)	(0.041)
	2,359	1,381	978
log(full-time employment)	-0.045^{*}	-0.038	-0.061^{*}
	(0.024)	(0.035)	(0.032)
	2,359	1,381	978
log(wage sum, soc. sec.)	-0.041	-0.060	-0.025
	(0.033)	(0.047)	(0.045)
	2,359	1,381	978
log(wage sum, survey)	-0.040	-0.124^{***}	0.051
	(0.033)	(0.044)	(0.050)
	2,268	1,329	939
Skill composition			
share of low-skilled	-0.006	-0.001	-0.012^{**}
	(0.005)	(0.009)	(0.005)
	2,359	1,381	978
share of medium-skilled	0.003	0.003	0.002
	(0.010)	(0.011)	(0.018)
	2,359	1,381	978
share of high-skilled	0.001	-0.003	0.007
	(0.009)	(0.007)	(0.017)
	2,359	1,381	978

This table shows the results from the fixed-effects difference-in-differences estimation with 1996 to 2000 as pre-treatment period and 2001 to 2005 as the treatment period. The estimation equation is: $y_{i,t} = \beta_0 + \beta_1 \,\mathsf{DSL\ access}_i\, D_{post,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{post,t}$ is an indicator variable for the treatment period from 2001 on. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 2 and 3.5 km from the MDF and zero for firms located between 3.5 and 5 km from the MDF. The table reports the results for β_1. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.2.7 Robustness checks speed upgrades: 0.5-2 km vs. 2-3.5 km

Table A.17: Results on speed upgrades, 0.5-2 km to 2-3.5 km: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	−0.014	0.021	0.018	0.035
	(0.087)	(0.030)	(0.083)	(0.039)
$D_{DSL} \times D_{2002}$	0.022	−0.002	0.037	0.005
	(0.089)	(0.030)	(0.084)	(0.041)
$D_{DSL} \times D_{2003}$	0.082	0.010	0.107	0.040
	(0.078)	(0.031)	(0.074)	(0.042)
$D_{DSL} \times D_{2004}$	0.109	0.007	0.156*	0.090**
	(0.086)	(0.032)	(0.083)	(0.043)
$D_{DSL} \times D_{2005}$	0.176*	0.019	0.216**	0.077*
	(0.106)	(0.035)	(0.098)	(0.044)
$D_{DSL} \times D_{2006}$	0.145	−0.007	0.165	0.005
	(0.106)	(0.033)	(0.101)	(0.043)
$D_{DSL} \times D_{2007}$	0.106	−0.045	0.120	0.011
	(0.108)	(0.032)	(0.102)	(0.045)
$D_{DSL} \times D_{2008}$	0.043	−0.043	0.065	0.010
	(0.089)	(0.036)	(0.086)	(0.048)
$D_{DSL} \times D_{2009}$	0.035	−0.050	0.068	0.012
	(0.087)	(0.038)	(0.083)	(0.052)
$D_{DSL} \times D_{2010}$	0.041	−0.020	0.054	0.023
	(0.089)	(0.039)	(0.085)	(0.051)
$D_{DSL} \times D_{2011}$	0.019	0.011	0.058	0.018
	(0.107)	(0.044)	(0.104)	(0.052)
R-squared	0.947	0.979	0.958	0.974
Obs.	7854	7854	7854	7605

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2 and 3.5 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data.

Table A.18: Results on speed upgrades, 0.5-2 km to 2-3.5 km: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.013	0.073	−0.009	0.000	−0.016
	(0.036)	(0.066)	(0.018)	(0.100)	(0.043)
$D_{DSL} \times D_{2002}$	0.011	−0.048	0.000	−0.028	−0.024
	(0.035)	(0.066)	(0.018)	(0.108)	(0.045)
$D_{DSL} \times D_{2003}$	0.065*	0.057	−0.001	−0.041	0.072
	(0.036)	(0.068)	(0.019)	(0.094)	(0.045)
$D_{DSL} \times D_{2004}$	0.104***	0.052	0.013	−0.027	0.103**
	(0.037)	(0.064)	(0.019)	(0.102)	(0.046)
$D_{DSL} \times D_{2005}$	0.050	−0.024	−0.007	−0.193	0.059
	(0.039)	(0.066)	(0.019)	(0.122)	(0.048)
$D_{DSL} \times D_{2006}$	0.061	0.084	−0.020	−0.103	0.067
	(0.039)	(0.063)	(0.019)	(0.122)	(0.050)
$D_{DSL} \times D_{2007}$	0.057	0.063	−0.018	−0.104	0.073
	(0.041)	(0.069)	(0.020)	(0.121)	(0.051)
$D_{DSL} \times D_{2008}$	0.059	0.070	−0.012	−0.010	0.124**
	(0.042)	(0.071)	(0.020)	(0.100)	(0.050)
$D_{DSL} \times D_{2009}$	0.058	0.074	−0.000	0.001	0.154***
	(0.041)	(0.072)	(0.023)	(0.101)	(0.051)
$D_{DSL} \times D_{2010}$	0.080*	0.032	0.004	−0.019	0.107**
	(0.041)	(0.074)	(0.023)	(0.105)	(0.051)
$D_{DSL} \times D_{2011}$	0.068	0.116	−0.023	0.013	0.132**
	(0.044)	(0.078)	(0.024)	(0.104)	(0.052)
R-squared	0.977	0.950	0.686	0.922	0.829
Obs.	7854	6334	6503	6859	6842

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2 and 3.5 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data.

A.2.8 Other robustness checks

Table A.19: Founded 1992 or later: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	−0.042	0.026	−0.004	−0.009
	(0.074)	(0.048)	(0.074)	(0.067)
$D_{DSL} \times D_{2002}$	0.032	0.020	0.044	−0.016
	(0.079)	(0.048)	(0.077)	(0.068)
$D_{DSL} \times D_{2003}$	−0.038	0.011	−0.000	−0.007
	(0.083)	(0.052)	(0.080)	(0.071)
$D_{DSL} \times D_{2004}$	0.044	0.045	0.096	0.096
	(0.086)	(0.060)	(0.086)	(0.079)
$D_{DSL} \times D_{2005}$	0.028	0.075	0.110	0.131
	(0.084)	(0.062)	(0.086)	(0.085)
$D_{DSL} \times D_{2006}$	0.081	0.052	0.137*	0.067
	(0.080)	(0.054)	(0.081)	(0.075)
$D_{DSL} \times D_{2007}$	0.094	0.043	0.118	0.109
	(0.079)	(0.053)	(0.079)	(0.076)
$D_{DSL} \times D_{2008}$	0.085	0.086	0.115	0.200**
	(0.089)	(0.057)	(0.087)	(0.083)
$D_{DSL} \times D_{2009}$	0.121	0.077	0.169**	0.182**
	(0.084)	(0.058)	(0.086)	(0.088)
$D_{DSL} \times D_{2010}$	0.142	0.084	0.147	0.098
	(0.089)	(0.062)	(0.090)	(0.086)
$D_{DSL} \times D_{2011}$	0.132	0.112	0.178	0.081
	(0.118)	(0.073)	(0.119)	(0.085)
R-squared	0.941	0.972	0.951	0.965
Obs.	4,182	4,182	4,182	4,046

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\,D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data. The sample only contains firms founded 1992 or later.

Table A.20: Founded 1992 or later: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.012	0.176*	−0.038	−0.000	−0.002
	(0.057)	(0.098)	(0.027)	(0.078)	(0.065)
$D_{DSL} \times D_{2002}$	0.011	0.191**	−0.046*	−0.027	0.001
	(0.057)	(0.095)	(0.025)	(0.086)	(0.069)
$D_{DSL} \times D_{2003}$	0.070	0.169	−0.032	0.108	0.064
	(0.059)	(0.109)	(0.027)	(0.090)	(0.070)
$D_{DSL} \times D_{2004}$	0.099	0.257***	−0.050*	0.034	0.094
	(0.060)	(0.099)	(0.029)	(0.085)	(0.070)
$D_{DSL} \times D_{2005}$	0.061	0.194*	−0.065**	−0.013	0.045
	(0.062)	(0.104)	(0.028)	(0.087)	(0.074)
$D_{DSL} \times D_{2006}$	0.099	0.357***	−0.098***	0.006	0.056
	(0.064)	(0.107)	(0.032)	(0.089)	(0.077)
$D_{DSL} \times D_{2007}$	0.170**	0.291***	−0.022	0.024	0.157*
	(0.069)	(0.111)	(0.035)	(0.088)	(0.084)
$D_{DSL} \times D_{2008}$	0.152**	0.245**	−0.024	0.038	0.105
	(0.069)	(0.122)	(0.036)	(0.083)	(0.077)
$D_{DSL} \times D_{2009}$	0.131*	0.257**	−0.027	−0.053	0.115
	(0.072)	(0.117)	(0.038)	(0.088)	(0.075)
$D_{DSL} \times D_{2010}$	0.189***	0.331***	−0.028	0.032	0.163**
	(0.072)	(0.116)	(0.039)	(0.099)	(0.078)
$D_{DSL} \times D_{2011}$	0.210***	0.345***	−0.061*	0.068	0.186**
	(0.073)	(0.115)	(0.035)	(0.111)	(0.090)
R-squared	0.973	0.940	0.713	0.916	0.830
Obs.	4,182	3,378	3,468	3,619	3,608

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. The sample only contains firms founded 1992 or later.

Table A.21: Founded 1992 or before: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	0.008	0.098*	0.033	0.108
	(0.060)	(0.059)	(0.060)	(0.082)
$D_{DSL} \times D_{2002}$	0.053	0.098*	0.058	0.112
	(0.065)	(0.059)	(0.063)	(0.080)
$D_{DSL} \times D_{2003}$	0.069	0.117*	0.083	0.140*
	(0.063)	(0.062)	(0.061)	(0.084)
$D_{DSL} \times D_{2004}$	0.050	0.093	0.080	0.169*
	(0.069)	(0.065)	(0.069)	(0.088)
$D_{DSL} \times D_{2005}$	0.042	0.130*	0.095	0.157*
	(0.067)	(0.067)	(0.068)	(0.091)
$D_{DSL} \times D_{2006}$	0.032	0.122*	0.047	0.100
	(0.067)	(0.067)	(0.066)	(0.093)
$D_{DSL} \times D_{2007}$	0.055	0.110*	0.087	0.105
	(0.065)	(0.066)	(0.066)	(0.097)
$D_{DSL} \times D_{2008}$	0.097	0.140**	0.105	0.187*
	(0.071)	(0.069)	(0.070)	(0.099)
$D_{DSL} \times D_{2009}$	0.064	0.150**	0.087	0.199**
	(0.071)	(0.071)	(0.067)	(0.098)
$D_{DSL} \times D_{2010}$	0.095	0.129*	0.099	0.159*
	(0.074)	(0.069)	(0.075)	(0.091)
$D_{DSL} \times D_{2011}$	0.198**	0.176**	0.230**	0.137
	(0.098)	(0.075)	(0.100)	(0.094)
R-squared	0.962	0.984	0.969	0.978
Obs.	4,062	4,062	4,062	3,932

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Information on total employment and monthly wage sum (last column) retrieved from survey data. The sample only contains firms founded 1992 or earlier.

Table A.22: Founded 1992 or before: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.010	0.154	−0.004	0.010	−0.036
	(0.067)	(0.108)	(0.032)	(0.070)	(0.058)
$D_{DSL} \times D_{2002}$	0.056	0.103	0.014	0.003	−0.028
	(0.061)	(0.102)	(0.028)	(0.081)	(0.065)
$D_{DSL} \times D_{2003}$	0.150**	0.250**	0.004	0.095	0.069
	(0.060)	(0.125)	(0.035)	(0.083)	(0.062)
$D_{DSL} \times D_{2004}$	0.196***	0.297**	−0.006	0.111	0.071
	(0.068)	(0.121)	(0.034)	(0.083)	(0.066)
$D_{DSL} \times D_{2005}$	0.160**	0.236**	−0.023	0.107	0.082
	(0.070)	(0.116)	(0.031)	(0.092)	(0.077)
$D_{DSL} \times D_{2006}$	0.154**	0.333***	−0.047	0.136	0.052
	(0.072)	(0.121)	(0.037)	(0.092)	(0.077)
$D_{DSL} \times D_{2007}$	0.203***	0.378***	−0.020	0.111	0.085
	(0.078)	(0.130)	(0.043)	(0.089)	(0.081)
$D_{DSL} \times D_{2008}$	0.152**	0.207	0.026	0.040	0.034
	(0.070)	(0.140)	(0.044)	(0.087)	(0.073)
$D_{DSL} \times D_{2009}$	0.211***	0.304**	0.010	0.156+	0.142*
	(0.070)	(0.130)	(0.044)	(0.087)	(0.072)
$D_{DSL} \times D_{2010}$	0.173**	0.240*	0.003	0.046	0.042
	(0.070)	(0.131)	(0.047)	(0.084)	(0.068)
$D_{DSL} \times D_{2011}$	0.167**	0.281**	−0.015	−0.018	0.082
	(0.072)	(0.117)	(0.034)	(0.111)	(0.078)
R-squared	0.975	0.948	0.699	0.939	0.848
Observations	4,062	3,272	3,354	3,554	3,546

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\,D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Robust standard errors in parentheses. Constant and year dummies included but now shown. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. The sample only contains firms founded 1992 or before.

Table A.23: West Germany: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	0.040	0.196	0.071	0.248
	(0.069)	(0.129)	(0.074)	(0.157)
$D_{DSL} \times D_{2002}$	0.046	0.191	0.030	0.288*
	(0.080)	(0.134)	(0.091)	(0.174)
$D_{DSL} \times D_{2003}$	0.127	0.209	0.141	0.292*
	(0.086)	(0.140)	(0.095)	(0.176)
$D_{DSL} \times D_{2004}$	0.105	0.288*	0.132	0.463**
	(0.084)	(0.147)	(0.099)	(0.189)
$D_{DSL} \times D_{2005}$	0.117	0.263*	0.132	0.324**
	(0.087)	(0.148)	(0.093)	(0.164)
$D_{DSL} \times D_{2006}$	0.157*	0.276*	0.079	0.236
	(0.090)	(0.152)	(0.100)	(0.182)
$D_{DSL} \times D_{2007}$	0.098	0.204	0.073	0.215
	(0.100)	(0.154)	(0.114)	(0.193)
$D_{DSL} \times D_{2008}$	0.064	0.145	0.038	0.219
	(0.105)	(0.155)	(0.115)	(0.193)
$D_{DSL} \times D_{2009}$	0.025	0.139	0.066	0.254
	(0.104)	(0.156)	(0.104)	(0.206)
$D_{DSL} \times D_{2010}$	−0.076	0.105	−0.085	0.149
	(0.101)	(0.142)	(0.121)	(0.173)
$D_{DSL} \times D_{2011}$	−0.068	0.140	−0.052	0.110
	(0.138)	(0.144)	(0.133)	(0.169)
R-squared	0.965	0.984	0.969	0.975
Obs.	2,600	2,600	2,600	2,495

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data. The sample only contains firms in West Germany.

Table A.24: West Germany: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.001	0.008	0.096**	0.081	−0.149
	(0.119)	(0.174)	(0.041)	(0.139)	(0.110)
$D_{DSL} \times D_{2002}$	0.068	0.067	0.019	0.019	−0.149
	(0.120)	(0.192)	(0.041)	(0.158)	(0.118)
$D_{DSL} \times D_{2003}$	0.227*	0.343	0.004	0.108	−0.023
	(0.122)	(0.233)	(0.042)	(0.167)	(0.113)
$D_{DSL} \times D_{2004}$	0.272**	0.421*	0.003	0.207	0.009
	(0.135)	(0.216)	(0.041)	(0.156)	(0.114)
$D_{DSL} \times D_{2005}$	0.343**	0.360*	−0.005	0.257	0.043
	(0.143)	(0.212)	(0.040)	(0.187)	(0.152)
$D_{DSL} \times D_{2006}$	0.328**	0.332	0.000	0.212	0.072
	(0.143)	(0.204)	(0.039)	(0.187)	(0.144)
$D_{DSL} \times D_{2007}$	0.305**	0.561**	−0.074*	0.208	−0.103
	(0.139)	(0.222)	(0.043)	(0.172)	(0.116)
$D_{DSL} \times D_{2008}$	0.251**	0.284	0.014	0.196	−0.010
	(0.127)	(0.208)	(0.039)	(0.181)	(0.151)
$D_{DSL} \times D_{2009}$	0.289**	0.280	0.008	0.342**	0.224*
	(0.127)	(0.190)	(0.038)	(0.171)	(0.133)
$D_{DSL} \times D_{2010}$	0.214*	0.263	−0.009	0.261*	0.080
	(0.127)	(0.197)	(0.040)	(0.154)	(0.122)
$D_{DSL} \times D_{2011}$	0.178	0.288	−0.030	0.248	0.114
	(0.129)	(0.195)	(0.039)	(0.180)	(0.127)
R-squared	0.968	0.940	0.679	0.928	0.781
Obs.	2,600	2,056	2,123	2,274	2,267

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. The sample only contains firms in West Germany.

Table A.25: Excluding municipalities without broadband internet access: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	−0.009	0.082	0.026	0.090
	(0.056)	(0.051)	(0.057)	(0.068)
$D_{DSL} \times D_{2002}$	0.048	0.076	0.054	0.080
	(0.060)	(0.051)	(0.059)	(0.068)
$D_{DSL} \times D_{2003}$	0.013	0.077	0.047	0.098
	(0.063)	(0.054)	(0.062)	(0.072)
$D_{DSL} \times D_{2004}$	0.051	0.112*	0.100	0.208***
	(0.066)	(0.061)	(0.066)	(0.077)
$D_{DSL} \times D_{2005}$	0.052	0.124**	0.119*	0.189**
	(0.064)	(0.060)	(0.066)	(0.077)
$D_{DSL} \times D_{2006}$	0.078	0.100*	0.107*	0.099
	(0.062)	(0.058)	(0.064)	(0.073)
$D_{DSL} \times D_{2007}$	0.064	0.077	0.089	0.137*
	(0.062)	(0.057)	(0.063)	(0.078)
$D_{DSL} \times D_{2008}$	0.052	0.101*	0.068	0.171**
	(0.069)	(0.060)	(0.069)	(0.082)
$D_{DSL} \times D_{2009}$	0.058	0.086	0.105	0.188**
	(0.066)	(0.062)	(0.067)	(0.086)
$D_{DSL} \times D_{2010}$	0.061	0.076	0.065	0.094
	(0.071)	(0.060)	(0.074)	(0.079)
$D_{DSL} \times D_{2011}$	0.074	0.114*	0.113	0.095
	(0.091)	(0.068)	(0.091)	(0.078)
R-squared	0.951	0.978	0.960	0.971
Obs.	5,956	5,956	5,956	5,760

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data. Firms in municipalities where no household has DSL access in 2005 or later years are excluded. Information on the share of household with DSL access retrieved from Federal Ministry of Economics and Technology (2009).

Table A.26: Excluding municipalities without broadband internet access: performance

	log revenues	log(VA)	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.007	0.109	0.001	0.008	−0.043
	(0.057)	(0.091)	(0.026)	(0.069)	(0.058)
$D_{DSL} \times D_{2002}$	0.029	0.111	−0.011	−0.016	−0.049
	(0.054)	(0.091)	(0.024)	(0.076)	(0.064)
$D_{DSL} \times D_{2003}$	0.123**	0.190*	−0.004	0.120	0.058
	(0.055)	(0.104)	(0.026)	(0.082)	(0.063)
$D_{DSL} \times D_{2004}$	0.158***	0.288***	−0.026	0.107	0.076
	(0.060)	(0.099)	(0.027)	(0.077)	(0.063)
$D_{DSL} \times D_{2005}$	0.122*	0.201**	−0.035	0.054	0.031
	(0.062)	(0.100)	(0.026)	(0.083)	(0.070)
$D_{DSL} \times D_{2006}$	0.125**	0.262**	−0.041	0.051	0.037
	(0.063)	(0.102)	(0.031)	(0.084)	(0.071)
$D_{DSL} \times D_{2007}$	0.166**	0.253**	−0.005	0.072	0.051
	(0.066)	(0.108)	(0.033)	(0.081)	(0.072)
$D_{DSL} \times D_{2008}$	0.127**	0.108	0.019	0.077	0.031
	(0.063)	(0.104)	(0.030)	(0.079)	(0.071)
$D_{DSL} \times D_{2009}$	0.155**	0.208**	−0.002	0.078	0.116*
	(0.065)	(0.105)	(0.034)	(0.083)	(0.070)
$D_{DSL} \times D_{2010}$	0.156**	0.229**	−0.004	0.075	0.096
	(0.065)	(0.105)	(0.033)	(0.083)	(0.067)
$D_{DSL} \times D_{2011}$	0.159**	0.257**	−0.035	0.079	0.122*
	(0.067)	(0.103)	(0.030)	(0.093)	(0.074)
R-squared	0.973	0.944	0.703	0.924	0.823
Obs.	5,956	4,799	4,925	5,194	5,178

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. Firms in municipalities where no household has DSL access in 2005 or later years are excluded. Information on the share of household with DSL access retrieved from Federal Ministry of Economics and Technology (2009).

Table A.27: Excluding counties with VDSL till 2008: firm size

	log(employment)		log(wage sum)	
	full-time	total	daily	survey
$D_{DSL} \times D_{2001}$	−0.063	0.073	−0.035	0.056
	(0.050)	(0.051)	(0.044)	(0.079)
$D_{DSL} \times D_{2002}$	0.055	0.083	0.055	0.052
	(0.075)	(0.060)	(0.071)	(0.087)
$D_{DSL} \times D_{2003}$	0.009	0.076	0.050	0.085
	(0.079)	(0.068)	(0.072)	(0.088)
$D_{DSL} \times D_{2004}$	0.051	0.128	0.110	0.236**
	(0.087)	(0.078)	(0.087)	(0.106)
$D_{DSL} \times D_{2005}$	0.086	0.157*	0.157*	0.222**
	(0.089)	(0.080)	(0.094)	(0.103)
$D_{DSL} \times D_{2006}$	0.074	0.144*	0.128	0.170*
	(0.090)	(0.084)	(0.091)	(0.103)
$D_{DSL} \times D_{2007}$	0.112	0.149*	0.154*	0.178*
	(0.089)	(0.083)	(0.089)	(0.098)
$D_{DSL} \times D_{2008}$	0.096	0.159*	0.130	0.228**
	(0.092)	(0.087)	(0.091)	(0.110)
$D_{DSL} \times D_{2009}$	0.062	0.133	0.105	0.200*
	(0.095)	(0.091)	(0.099)	(0.120)
$D_{DSL} \times D_{2010}$	0.060	0.135	0.061	0.125
	(0.101)	(0.092)	(0.105)	(0.117)
$D_{DSL} \times D_{2011}$	0.019	0.158	0.054	0.109
	(0.124)	(0.102)	(0.128)	(0.114)
R-squared	0.943	0.972	0.952	0.965
Obs.	4,349	4,349	4,349	4,177

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on total employment and monthly wage sum (last column) retrieved from survey data. The sample contains non-manufacturing firms in counties in which VDSL was not introduced till 2008.

Table A.28: Excluding counties with VDSL till 2008: performance

	log revenues	log VA	Shr. of Inputs	log(rev. p. FTE)	log(rev. p. empl.)
$D_{DSL} \times D_{2001}$	0.003	0.107	−0.001	0.039	−0.026
	(0.037)	(0.087)	(0.022)	(0.071)	(0.061)
$D_{DSL} \times D_{2002}$	0.049	0.171*	−0.012	−0.007	−0.010
	(0.063)	(0.101)	(0.027)	(0.097)	(0.069)
$D_{DSL} \times D_{2003}$	0.123*	0.177	0.002	0.094	0.024
	(0.074)	(0.126)	(0.026)	(0.095)	(0.074)
$D_{DSL} \times D_{2004}$	0.171**	0.241**	−0.007	0.084	0.047
	(0.081)	(0.123)	(0.026)	(0.104)	(0.077)
$D_{DSL} \times D_{2005}$	0.169*	0.222*	−0.022	0.053	0.053
	(0.093)	(0.128)	(0.028)	(0.122)	(0.095)
$D_{DSL} \times D_{2006}$	0.175*	0.306**	−0.046	0.077	0.031
	(0.099)	(0.137)	(0.028)	(0.125)	(0.099)
$D_{DSL} \times D_{2007}$	0.194*	0.299*	−0.024	−0.002	−0.028
	(0.104)	(0.156)	(0.044)	(0.113)	(0.086)
$D_{DSL} \times D_{2008}$	0.115	0.161	0.009	−0.017	−0.068
	(0.096)	(0.153)	(0.042)	(0.099)	(0.088)
$D_{DSL} \times D_{2009}$	0.112	0.208	0.002	0.018	0.021
	(0.095)	(0.157)	(0.045)	(0.111)	(0.084)
$D_{DSL} \times D_{2010}$	0.151	0.261*	−0.016	0.084	0.060
	(0.102)	(0.150)	(0.046)	(0.114)	(0.090)
$D_{DSL} \times D_{2011}$	0.162	0.235	−0.033	0.136	0.074
	(0.101)	(0.146)	(0.040)	(0.131)	(0.101)
R-squared	0.969	0.936	0.703	0.905	0.814
Obs.	4,349	3,526	3,631	3,789	3,776

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Information on revenues, value added, share of inputs, and total employment retrieved from survey data. The number of full-time employees (FTE) is calculated based on social-security data. The sample contains non-manufacturing firms in counties in which VDSL was not introduced till 2008.

A.2.9 Other outcome variables

Table A.29: Outcome variable: firms exits

	Non-manuf.	Manuf.	Pooled
$D_{DSL} \times D_{2001}$	−0.002	−0.000	−0.001
	(0.002)	(0.000)	(0.001)
$D_{DSL} \times D_{2002}$	−0.006*	−0.001	−0.004**
	(0.003)	(0.001)	(0.002)
$D_{DSL} \times D_{2003}$	0.002	−0.004	0.000
	(0.006)	(0.003)	(0.004)
$D_{DSL} \times D_{2004}$	−0.007	−0.009	−0.008
	(0.020)	(0.014)	(0.013)
$D_{DSL} \times D_{2005}$	−0.016	−0.002	−0.009
	(0.024)	(0.013)	(0.015)
$D_{DSL} \times D_{2006}$	−0.013	−0.009	−0.011
	(0.024)	(0.017)	(0.015)
$D_{DSL} \times D_{2007}$	−0.012	0.001	−0.007
	(0.027)	(0.015)	(0.017)
$D_{DSL} \times D_{2008}$	−0.022	−0.000	−0.014
	(0.030)	(0.016)	(0.019)
$D_{DSL} \times D_{2009}$	−0.020	0.001	−0.012
	(0.030)	(0.017)	(0.019)
$D_{DSL} \times D_{2010}$	0.000	0.005	0.002
	(0.030)	(0.018)	(0.019)
$D_{DSL} \times D_{2011}$	−0.014	0.020	−0.000
	(0.030)	(0.020)	(0.019)
R-squared	0.511	0.405	0.484
Observations	8,720	6,501	15,275

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.30: Outcome variable: firm invests in ICT, panel 1996-2005

	Pooled	Non-manufacturing	Manufacturing
$D_{DSL} \times D_{post}$	-0.024	0.037	0.017
	(0.047)	(0.052)	(0.077)
R-squared	0.489	0.450	0.456
Observations	2,027	1,255	929

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{post} + \alpha_i + \alpha_t + u_{it}$, where D_{post} is an indicator variable for the post-treatment period. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Investments in ICT retrieved from survey data. Firms are asked whether they invested in ICT. Samples are split by firms reporting to invest in ICT in 2000 or not.

Table A.31: Outcome variable: firm invests in ICT, panel 2000-2011

	Pooled	Non-manufacturing	Manufacturing
$D_{DSL} \times D_{2001}$	0.031	−0.062	−0.031
	(0.038)	(0.057)	(0.049)
$D_{DSL} \times D_{2002}$	−0.003	−0.076	0.014
	(0.040)	(0.058)	(0.050)
$D_{DSL} \times D_{2003}$	−0.059	0.075	0.037
	(0.045)	(0.064)	(0.055)
$D_{DSL} \times D_{2004}$	−0.053	0.072	0.060
	(0.045)	(0.066)	(0.056)
$D_{DSL} \times D_{2005}$	−0.011	−0.014	0.024
	(0.049)	(0.068)	(0.062)
$D_{DSL} \times D_{2006}$	−0.066	0.002	0.102*
	(0.047)	(0.068)	(0.061)
$D_{DSL} \times D_{2007}$	−0.042	0.008	0.061
	(0.049)	(0.068)	(0.066)
$D_{DSL} \times D_{2008}$	0.050	−0.038	−0.092
	(0.053)	(0.073)	(0.069)
$D_{DSL} \times D_{2009}$	−0.036	0.018	0.039
	(0.052)	(0.075)	(0.067)
$D_{DSL} \times D_{2010}$	0.051	−0.042	−0.098
	(0.055)	(0.073)	(0.072)
$D_{DSL} \times D_{2011}$	−0.043	−0.041	0.091
	(0.061)	(0.080)	(0.083)
R-squared	0.506	0.463	0.457
Observations	11,345	6,274	6,117

This table shows the results from the fixed-effects difference-in-differences estimation with 2000 to 2002 as pre-treatment period, small speed upgrades in 2003 to 2005 and a major upgrade to ADSL2+ in 2006. The estimation equation is: $y_{i,t} = \beta_0 + \beta_{1,t}\,\mathsf{DSL\ access}_i\, D_{year,t} + \alpha_i + \alpha_t + u_{it}$, where $D_{year,t}$ is an indicator variable for each year. $\mathsf{DSL\ access}_i$ is an indicator variable equal to one if the firm is located between 0.5 and 2 km from the MDF and zero for firms located between 2.8 and 4.2 km from the MDF. The table reports the results for $\beta_{1,t}$. To adjust for different numbers of firms in the treatment and control groups by three-digit sector, I use the weights suggested by Iacus et al. (2009). Number of observations reported for each regression. Standard errors clustered at the firm-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Investments in ICT retrieved from survey data. Firms are asked whether they invested in ICT. Samples are split by firms reporting to invest in ICT in 2000 or not.

Appendix B

Appendix to Chapter 2

B.1 Data

B.1.1 Data sources and record linkage procedure

Social Security records

Employee history. The Integrated Employment Biographies (*Integrierte Erwerbsbiografien, IEB*) are based on records from the German Social Security System. They contain information on all employees subject to social insurance contributions since 1975 and are updated at least annually. The data cover nearly all private sector employees in Germany, but do not cover civil servants and the self-employed. The IEB contain information on birth year, gender, nationality, education, occupation, full time or part-time status and daily earnings of each employee. Jacobebbinghaus and Seth (2007) and Antoni et al. (2016) provide a detailed description of the structure of the data.[1]

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

Establishment history panel. The Establishment History Panel (*Betriebshistorikpanel, BHP*) is a panel data set that contains information on the sector, number of

[1] The paper by Antoni et al. (2016) focuses on the Sample of Integrated Labor Market Biographies (SIAB), a 2% random sample drawn from the IEB.

employees 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.[2] Sectors are defined based on the Classification of Economic Activities of the German Statistical Office (see also section C.1.2). The location of establishments is provided at the county level. Germany is divided into 402 counties with around 200,000 inhabitants on average. German counties are roughly comparable to counties in the US. Schmucker et al. (2016) provide a detailed description of the data set.

Extension file entry and exit. The extension file entry and exit uses information on worker flows to identify establishment openings and closings. Establishment identifiers may change when a firm restructures. The extension file helps mitigate bias related to restructurings. Hethey and Schmieder (2010) provide details on the file.

ORBIS

We use a linkage table between the Social Security Records and the firm-level database Orbis of the commercial data provider Bureau van Dijk (BvD). BvD compiles its firm-level data from publicly available sources as well as by acquiring data from other commercial data providers. For Germany, BvDs main data provider is Creditreform. Internationally, BvD offers more than 20 different databases with the main customers being privates companies looking for business intelligence on, for instance, competitors, business partners or potential targets for acquisitions. Within BvDs databases, a company is defined 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 at that time. In the years before these changes, a higher share of financial information is missing.

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

Appendix to Chapter 2

Record linkage procedure

The record linkage between Orbis and the Social Security data was performed independently of our project by the German Record Linkage Center (GRLC, see Antoni and Schnell (2017) or www.record-linkage.de for more details on the GRLC). The basis of the linkage was an extract of Orbis acquired by the Institute for Employment Research (IAB). This extract contained data on all German firms at the reference date of January 30, 2014. Of the 1,938,990 firms contained in the data, 1,627,668 were marked as active in Germany at that reference date.

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

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

As these identifiers are non-unique and error-prone, the GRLC developed extensive cleaning, standardization and parsing routines (usually referred to as pre-processing) to achieve records that could successfully be compared between the two data sources.

To deal with remaining differences in, for instance, the spelling or abbreviations of the identifiers, the GRLC applied error-tolerant methods of record linkage (see Christen, 2012). The resulting linkage process consists of 17 consecutive steps, not counting the pre-processing, that varied in terms of which identifiers were used and how strict the requirements on agreement of the compared records were. Schild (2016) provides a more detailed description of the record linkage process. Antoni et al. (2018) report on the linkage success and the representativeness of the resulting data set.

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

Identification of headquarters

The linkage process explained in the previous subsection aimed at identifying as many establishments per firm as possible without trying to determine which of the linked establishments had been the headquarters of the firm. This information was added by the Research Data Centre (FDZ) at the IAB afterwards. To do so, the FDZ performed several iterative steps that mainly relied on the address of the firm according to Orbis and of the establishments according to the administrative data. During later steps the FDZ also used information on the share of administrative staff or the industry code of the establishments under consideration. Given that the administrative data do not contain information that directly identifies an establishment as the headquarters of a superordinate company, this process had to rely on variables that allow the identification of the most likely headquarters among the linked establishments. Antoni et al. (2018) provide more details on the whole process and on the remaining uncertainty regarding the identified headquarters.

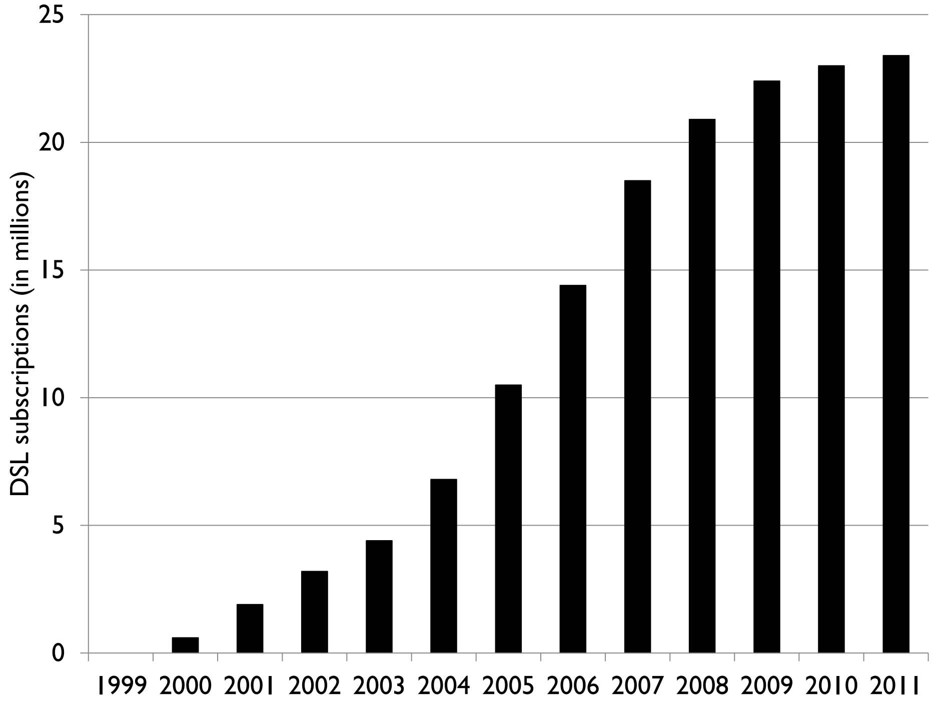
B.1.2 Sector and occupation classification

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

The information on the occupation of employees follows the German classification of occupations *"Klassifikation der Berufe"* (KldB). The years 1998-2010 contain the three digit occupation according to the 1988 version of the KldB.

B.1.3 Identification strategy

Figure B.1: Number of DSL subscriptions in Germany



The figure shows the number of DSL subscriptions in Germany over time Bundesnetzagentur (2011).

B.2 Results

B.2.1 Additional results

Table B.1: Results by initial skill composition: additional regressions

	Share of low-skilled Incl. skilled in 1999	Share of high-skilled Without skilled in 1999
HQ distance × $D_{2000-03}$	0.029	−0.027
	(0.087)	(0.047)
HQ distance × $D_{2004-05}$	0.032	−0.063
	(0.110)	(0.093)
HQ distance × $D_{2006-10}$	0.015	−0.112
	(0.143)	(0.109)
Sub-estab. distance × $D_{2000-03}$	−0.062	0.173***
	(0.086)	(0.048)
Sub-estab. distance × $D_{2004-05}$	−0.110	0.287**
	(0.112)	(0.091)
Sub-estab. distance × $D_{2006-10}$	−0.123	0.405***
	(0.142)	(0.111)
R-squared	0.872	0.578
Obs.	51,239	103,518

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree. Samples are split by zero or positive employment of high-skilled labor in 1999.

B.2.2 Yearly effects

Table B.2: Employment at subordinate establishment level: yearly effects

		Skill shares		
	Employment	Low	Medium	High
Sub-e. distance × D_{2000}	0.001	−0.120	0.172^{+}	−0.052
	(0.005)	(0.085)	(0.093)	(0.054)
Sub-e. distance × D_{2001}	−0.013*	−0.088	0.052	0.036
	(0.006)	(0.099)	(0.114)	(0.068)
Sub-e. distance × D_{2002}	−0.008	−0.194^{+}	0.150	0.044
	(0.007)	(0.109)	(0.137)	(0.095)
Sub-e. distance × D_{2003}	0.026***	−0.104	0.056	0.048
	(0.007)	(0.111)	(0.146)	(0.104)
Sub-e. distance × D_{2004}	0.019*	−0.253*	0.141	0.111
	(0.008)	(0.119)	(0.156)	(0.113)
Sub-e. distance × D_{2005}	0.016*	−0.343**	0.190	0.153
	(0.008)	(0.130)	(0.172)	(0.114)
Sub-e. distance × D_{2006}	0.019*	−0.306*	0.070	0.236*
	(0.008)	(0.126)	(0.167)	(0.114)
Sub-e. distance × D_{2007}	0.034***	−0.347**	0.023	0.325**
	(0.008)	(0.132)	(0.176)	(0.124)
Sub-e. distance × D_{2008}	0.032***	−0.411**	0.146	0.265^{+}
	(0.009)	(0.134)	(0.185)	(0.137)
Sub-e. distance × D_{2009}	0.024*	−0.429**	0.171	0.258*
	(0.010)	(0.142)	(0.185)	(0.125)
Sub-e. distance × D_{2010}	0.022*	−0.436**	0.187	0.249^{+}
	(0.009)	(0.134)	(0.182)	(0.131)
HQ distance × D_{2000}	−0.003	0.118	−0.192*	0.074
	(0.004)	(0.084)	(0.092)	(0.053)
HQ distance × D_{2001}	0.010^{+}	0.069	−0.070	0.001
	(0.006)	(0.097)	(0.109)	(0.065)
HQ distance × D_{2002}	0.001	0.149	−0.147	−0.001
	(0.007)	(0.107)	(0.133)	(0.092)
HQ distance × D_{2003}	−0.039***	0.041	−0.041	−0.000
	(0.007)	(0.110)	(0.144)	(0.101)
HQ distance × D_{2004}	−0.035***	0.163	−0.115	−0.048
	(0.008)	(0.118)	(0.153)	(0.109)
HQ distance × D_{2005}	−0.034***	0.250^{+}	−0.178	−0.072
	(0.008)	(0.130)	(0.168)	(0.108)
HQ distance × D_{2006}	−0.038***	0.199	−0.058	−0.140
	(0.008)	(0.125)	(0.164)	(0.109)
HQ distance × D_{2007}	−0.055***	0.231^{+}	−0.007	−0.224^{+}
	(0.008)	(0.132)	(0.173)	(0.116)
HQ distance × D_{2008}	−0.052***	0.290*	−0.164	−0.126
	(0.009)	(0.134)	(0.179)	(0.130)
HQ distance × D_{2009}	−0.046***	0.292*	−0.188	−0.104
	(0.009)	(0.142)	(0.182)	(0.122)
HQ distance × D_{2010}	−0.045***	0.269*	−0.179	−0.090
	(0.009)	(0.135)	(0.181)	(0.128)
R-squared	0.902	0.759	0.828	0.884
Obs.	217,387	217,387	217,387	217,387

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \mathsf{log(HQ\ distance\ to\ MDF)}_i \times Year_t + \beta_{2,t} \times \mathsf{log(subordinate\ establishment\ distance\ to\ MDF)}_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. $\mathsf{log(HQ\ distance\ to\ MDF)}_i$ and $\mathsf{log(subordinate\ establishment\ distance\ to\ MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Table B.3: Employment at HQ level: yearly effects

		Skill shares		
	Employment	Low	Medium	High
HQ distance × D_{2000}	0.007**	−0.042	0.031	0.011
	(0.003)	(0.033)	(0.045)	(0.037)
HQ distance × D_{2001}	0.005	−0.063*	0.084	−0.021
	(0.003)	(0.034)	(0.052)	(0.044)
HQ distance × D_{2002}	0.006*	−0.109***	0.043	0.066
	(0.003)	(0.040)	(0.064)	(0.055)
HQ distance × D_{2003}	0.003	−0.176***	0.105*	0.071
	(0.004)	(0.041)	(0.061)	(0.055)
HQ distance × D_{2004}	0.001	−0.203***	0.047	0.155***
	(0.004)	(0.044)	(0.061)	(0.054)
HQ distance × D_{2005}	−0.000	−0.190***	0.002	0.188***
	(0.004)	(0.051)	(0.071)	(0.061)
HQ distance × D_{2006}	0.002	−0.178***	0.011	0.167***
	(0.005)	(0.055)	(0.076)	(0.062)
HQ distance × D_{2007}	0.005	−0.158***	−0.029	0.186***
	(0.005)	(0.055)	(0.076)	(0.063)
HQ distance × D_{2008}	0.009*	−0.215***	−0.006	0.221***
	(0.005)	(0.059)	(0.082)	(0.068)
HQ distance × D_{2009}	0.008	−0.241***	−0.022	0.264***
	(0.005)	(0.058)	(0.082)	(0.070)
HQ distance × D_{2010}	0.008	−0.254***	−0.043	0.297***
	(0.005)	(0.059)	(0.088)	(0.077)
Sub-e. distance × D_{2000}	−0.003	0.026	−0.032	0.006
	(0.004)	(0.035)	(0.049)	(0.041)
Sub-e. distance × D_{2001}	0.001	0.039	−0.074	0.035
	(0.004)	(0.037)	(0.057)	(0.050)
Sub-e. distance × D_{2002}	−0.001	0.060	−0.047	−0.013
	(0.004)	(0.041)	(0.067)	(0.057)
Sub-e. distance × D_{2003}	0.001	0.093**	−0.071	−0.022
	(0.004)	(0.040)	(0.064)	(0.061)
Sub-e. distance × D_{2004}	0.001	0.079*	−0.031	−0.048
	(0.004)	(0.042)	(0.062)	(0.058)
Sub-e. distance × D_{2005}	0.000	0.052	0.018	−0.070
	(0.005)	(0.052)	(0.070)	(0.061)
Sub-e. distance × D_{2006}	−0.001	0.016	0.013	−0.029
	(0.005)	(0.049)	(0.068)	(0.056)
Sub-e. distance × D_{2007}	−0.001	−0.015	0.025	−0.010
	(0.005)	(0.052)	(0.074)	(0.063)
Sub-e. distance × D_{2008}	−0.004	0.029	−0.023	−0.006
	(0.005)	(0.056)	(0.078)	(0.068)
Sub-e. distance × D_{2009}	−0.004	0.024	0.000	−0.025
	(0.005)	(0.054)	(0.080)	(0.071)
Sub-e. distance × D_{2010}	−0.004	0.015	−0.007	−0.008
	(0.005)	(0.058)	(0.089)	(0.080)
R-squared	0.959	0.866	0.888	0.919
Obs.	55,047	55,047	55,047	55,047

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Year_t + \beta_{2,t} \times \text{avg. subordinate establishment distance to MDF}_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the average subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.4: Employment at firm level: yearly effects

		Skill shares		
	Employment	Low	Medium	High
HQ distance × D_{2000}	0.006**	−0.003	−0.001	0.042
	(0.002)	(0.028)	(0.039)	(0.030)
HQ distance × D_{2001}	0.004	−0.035	0.050	0.038
	(0.003)	(0.033)	(0.048)	(0.033)
HQ distance × D_{2002}	0.005*	−0.083**	0.054	0.076**
	(0.003)	(0.034)	(0.053)	(0.035)
HQ distance × D_{2003}	0.002	−0.129***	0.132**	0.046
	(0.003)	(0.034)	(0.057)	(0.038)
HQ distance × D_{2004}	−0.000	−0.189***	0.146***	0.094**
	(0.003)	(0.040)	(0.056)	(0.036)
HQ distance × D_{2005}	−0.000	−0.188***	0.118*	0.118***
	(0.003)	(0.045)	(0.063)	(0.041)
HQ distance × D_{2006}	0.003	−0.193***	0.144*	0.108**
	(0.003)	(0.047)	(0.074)	(0.047)
HQ distance × D_{2007}	0.004	−0.167***	0.109	0.139***
	(0.004)	(0.046)	(0.075)	(0.049)
HQ distance × D_{2008}	0.008**	−0.212***	0.114	0.165***
	(0.004)	(0.048)	(0.080)	(0.053)
HQ distance × D_{2009}	0.005	−0.222***	0.104	0.190***
	(0.004)	(0.049)	(0.083)	(0.056)
HQ distance × D_{2010}	0.006	−0.268***	0.098	0.207***
	(0.004)	(0.051)	(0.088)	(0.057)
Sub-e. distance × D_{2000}	−0.002	−0.020	0.040	−0.035
	(0.003)	(0.031)	(0.045)	(0.033)
Sub-e. distance × D_{2001}	0.002	0.008	0.004	−0.032
	(0.003)	(0.037)	(0.053)	(0.037)
Sub-e. distance × D_{2002}	0.000	0.031	−0.023	−0.045
	(0.003)	(0.036)	(0.055)	(0.037)
Sub-e. distance × D_{2003}	0.002	0.054	−0.079	−0.021
	(0.003)	(0.038)	(0.057)	(0.041)
Sub-e. distance × D_{2004}	0.002	0.080*	−0.097*	−0.020
	(0.003)	(0.043)	(0.058)	(0.041)
Sub-e. distance × D_{2005}	0.001	0.077	−0.062	−0.039
	(0.003)	(0.049)	(0.062)	(0.040)
Sub-e. distance × D_{2006}	−0.001	0.059	−0.068	0.004
	(0.003)	(0.047)	(0.067)	(0.043)
Sub-e. distance × D_{2007}	0.001	0.023	−0.051	0.002
	(0.003)	(0.049)	(0.071)	(0.046)
Sub-e. distance × D_{2008}	−0.000	0.050	−0.079	0.008
	(0.004)	(0.051)	(0.076)	(0.054)
Sub-e. distance × D_{2009}	0.001	0.031	−0.079	0.010
	(0.004)	(0.052)	(0.081)	(0.058)
Sub-e. distance × D_{2010}	−0.000	0.057	−0.088	0.025
	(0.004)	(0.053)	(0.087)	(0.060)
R-squared	0.972	0.902	0.903	0.949
Obs.	55,151	55,151	55,151	55,151

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times$ log(HQ distance to MDF)$_i \times Year_t + \beta_{2,t} \times$ avg. subordinate establishment distance to MDF$_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. log(HQ distance to MDF)$_i$ and log(subordinate establishment distance to MDF)$_j$ describe the distance of the HQ or the average subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.2.3 Robustness checks

Table B.5: Robustness checks: employment

	Founded before 1992	Western Germany	Mun. had DSL in 2005	No VDSL till 2008
Sub-e. distance × $D_{2000-03}$	0.025***	0.002	0.002	−0.018
	(0.007)	(0.006)	(0.005)	(0.015)
Sub-e. distance × $D_{2004-05}$	0.026*	0.020*	0.011	0.017
	(0.011)	(0.008)	(0.007)	(0.019)
Sub-e. distance × $D_{2006-10}$	0.030*	0.033***	0.022**	0.048*
	(0.012)	(0.009)	(0.008)	(0.019)
HQ distance × $D_{2000-03}$	−0.036***	−0.005	−0.007	0.010
	(0.007)	(0.006)	(0.005)	(0.014)
HQ distance × $D_{2004-05}$	−0.048***	−0.027***	−0.025***	−0.034+
	(0.011)	(0.008)	(0.007)	(0.018)
HQ distance × $D_{2006-10}$	−0.057***	−0.039***	−0.039***	−0.073**
	(0.011)	(0.008)	(0.007)	(0.018)
R-squared	0.923	0.904	0.902	0.877
Obs.	86,483	153,612	214,695	43,359

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^+$ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001. The first column reports the results for subordinate establishments founded before 1992. The second sample only contains subordinate establishments in Western Germany. The third sample excludes municipalities in which no household had DSL in 2005. The fourth sample excludes counties in which VDSL was introduced until 2008.

Table B.6: Robustness checks: share of low-skilled employment

	Founded before 1992	Western Germany	Mun. had DSL in 2005	No VDSL till 2008
Sub-e. distance × $D_{2000-03}$	-0.088	-0.114	-0.132	0.005
	(0.109)	(0.101)	(0.081)	(0.209)
Sub-e. distance × $D_{2004-05}$	-0.355^{*}	-0.344^{*}	-0.325^{**}	-0.515^{+}
	(0.177)	(0.137)	(0.111)	(0.287)
Sub-e. distance × $D_{2006-10}$	-0.261	-0.443^{**}	-0.401^{***}	-0.403
	(0.168)	(0.140)	(0.114)	(0.254)
HQ distance × $D_{2000-03}$	0.042	0.094	0.104	-0.045
	(0.108)	(0.097)	(0.079)	(0.212)
HQ distance × $D_{2004-05}$	0.269	0.267^{*}	0.250^{*}	0.430
	(0.178)	(0.132)	(0.109)	(0.295)
HQ distance × $D_{2006-10}$	0.104	0.313^{*}	0.293^{*}	0.291
	(0.170)	(0.135)	(0.114)	(0.255)
R-squared	0.797	0.750	0.759	0.770
Obs.	86,483	153,612	214,695	43,359

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \log(\text{HQ distance to MDF})_i \times Period_t + \beta_{2,t} \times \log(\text{subordinate establishment distance to MDF})_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\log(\text{HQ distance to MDF})_i$ and $\log(\text{subordinate establishment distance to MDF})_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The first column reports the results for subordinate establishments founded before 1992. The second sample only contains subordinate establishments in Western Germany. The third sample excludes municipalities in which no household had DSL in 2005. The fourth sample excludes counties in which VDSL was introduced until 2008.

Table B.7: Robustness checks: share of medium-skilled employment

	Founded before 1992	Western Germany	Mun. had DSL in 2005	No VDSL till 2008
Sub-e. distance × $D_{2000-03}$	0.091	0.090	0.118	0.033
	(0.137)	(0.114)	(0.095)	(0.267)
Sub-e. distance × $D_{2004-05}$	0.187	0.236	0.217	0.446
	(0.217)	(0.170)	(0.144)	(0.421)
Sub-e. distance × $D_{2006-10}$	0.017	0.161	0.142	0.202
	(0.214)	(0.182)	(0.152)	(0.400)
HQ distance × $D_{2000-03}$	−0.083	−0.094	−0.125	−0.034
	(0.135)	(0.109)	(0.092)	(0.266)
HQ distance × $D_{2004-05}$	−0.197	−0.220	−0.203	−0.420
	(0.213)	(0.162)	(0.139)	(0.428)
HQ distance × $D_{2006-10}$	−0.030	−0.159	−0.137	−0.193
	(0.211)	(0.178)	(0.146)	(0.410)
R-squared	0.842	0.808	0.828	0.845
Obs.	86,483	153,612	214,695	43,359

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \log(\text{HQ distance to MDF})_i \times Period_t + \beta_{2,t} \times \log(\text{subordinate establishment distance to MDF})_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\log(\text{HQ distance to MDF})_i$ and $\log(\text{subordinate establishment distance to MDF})_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The first column reports the results for subordinate establishments founded before 1992. The second sample only contains subordinate establishments in Western Germany. The third sample excludes municipalities in which no household had DSL in 2005. The fourth sample excludes counties in which VDSL was introduced until 2008.

Table B.8: Robustness checks: share of high-skilled employment

	Founded before 1992	Western Germany	Mun. had DSL in 2005	No VDSL till 2008
Sub-e. distance × $D_{2000-03}$	−0.003	0.023	0.014	−0.038
	(0.085)	(0.072)	(0.064)	(0.207)
Sub-e. distance × $D_{2004-05}$	0.169	0.108	0.108	0.069
	(0.141)	(0.108)	(0.099)	(0.276)
Sub-e. distance × $D_{2006-10}$	0.245	0.282*	0.259*	0.201
	(0.151)	(0.114)	(0.108)	(0.267)
HQ distance × $D_{2000-03}$	0.042	−0.000	0.021	0.078
	(0.081)	(0.070)	(0.062)	(0.207)
HQ distance × $D_{2004-05}$	−0.072	−0.047	−0.048	−0.010
	(0.135)	(0.105)	(0.094)	(0.273)
HQ distance × $D_{2006-10}$	−0.074	−0.154	−0.156	−0.098
	(0.145)	(0.114)	(0.102)	(0.272)
R-squared	0.896	0.877	0.884	0.884
Obs.	86,483	153,612	214,695	43,359

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \log(\text{HQ distance to MDF})_i \times Period_t + \beta_{2,t} \times \log(\text{subordinate establishment distance to MDF})_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\log(\text{HQ distance to MDF})_i$ and $\log(\text{subordinate establishment distance to MDF})_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The first column reports the results for subordinate establishments founded before 1992. The second sample only contains subordinate establishments in Western Germany. The third sample excludes municipalities in which no household had DSL in 2005. The fourth sample excludes counties in which VDSL was introduced until 2008.

B.2.4 Sample: max. 4.2 km

Table B.9: Subordinate establishments: max. 4.2 km

		Skill shares		
	Employment	Low	Medium	High
Sub-estab. distance × $D_{2000-03}$	0.002	-0.150^{+}	0.148	0.003
	(0.006)	(0.091)	(0.106)	(0.073)
Sub-estab. distance × $D_{2004-05}$	0.020^{*}	-0.344^{**}	0.269^{+}	0.075
	(0.008)	(0.123)	(0.156)	(0.108)
Sub-estab. distance × $D_{2006-10}$	0.033^{***}	-0.406^{**}	0.204	0.202^{+}
	(0.009)	(0.129)	(0.165)	(0.119)
HQ distance × $D_{2000-03}$	−0.007	0.120	−0.154	−0.034
	(0.006)	(0.089)	(0.103)	(0.070)
HQ distance × $D_{2004-05}$	-0.034^{***}	0.267^{*}	-0.257^{+}	−0.010
	(0.008)	(0.121)	(0.150)	(0.102)
HQ distance × $D_{2006-10}$	-0.053^{***}	0.294^{*}	−0.202	−0.091
	(0.008)	(0.129)	(0.160)	(0.112)
R-squared	0.901	0.759	0.829	0.885
Obs.	195,832	195,832	195,832	195,832

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \log(\text{HQ distance to MDF})_i \times Year_t + \beta_{2,t} \times \text{avg. subordinate establishment distance to MDF}_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. $\log(\text{HQ distance to MDF})_i$ and $\log(\text{subordinate establishment distance to MDF})_j$ describe the distance of the HQ or the average subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample only contains subordinate establishments located up to 4.2 km from the MDF. The same restriction applies to their HQ.

Table B.10: Subordinate establishments by initial skill composition: max. 4.2 km

	With high-skilled empl. in 1999			Without high-skilled empl. in 1999		
		Skillshares			Skillshares	
	Employment (1)	Medium (2)	High (3)	Employment (4)	Low (5)	Medium (6)
Sub-estab. distance × $D_{2000-03}$	0.030**	−0.021	0.069	0.011+	−0.140	−0.011
	(0.011)	(0.228)	(0.232)	(0.006)	(0.118)	(0.130)
Sub-estab. distance × $D_{2004-05}$	0.041*	−0.188	0.322	0.013	−0.529**	0.285
	(0.020)	(0.342)	(0.348)	(0.010)	(0.191)	(0.209)
Sub-estab. distance × $D_{2006-10}$	0.062**	−0.298	0.401	0.016	−0.565**	0.245
	(0.022)	(0.352)	(0.344)	(0.011)	(0.204)	(0.235)
HQ distance × $D_{2000-03}$	−0.045***	0.277	−0.287	−0.010	0.121	−0.123
	(0.012)	(0.227)	(0.232)	(0.006)	(0.113)	(0.124)
HQ distance × $D_{2004-05}$	−0.074***	0.527	−0.572+	−0.019+	0.461*	−0.450*
	(0.020)	(0.341)	(0.345)	(0.010)	(0.188)	(0.203)
HQ distance × $D_{2006-10}$	−0.101***	0.615+	−0.603+	−0.027**	0.428*	−0.412+
	(0.022)	(0.345)	(0.336)	(0.010)	(0.204)	(0.231)
R-squared	0.905	0.877	0.895	0.877	0.757	0.726
Obs.	47,173	47,173	47,173	91,509	91,509	91,509

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree. Samples are split by zero or positive employment of high-skilled labor in 1999. The sample only contains subordinate establishments located up to 4.2 km from the MDF. The same restriction applies to their HQ.

B.2.5 Sample: Non-manufacturing

Table B.11: Subordinate establishments: non-manufacturing

		Skill shares		
	Employment	Low	Medium	High
Sub-estab. distance × $D_{2000-03}$	0.004	−0.120	0.037	0.083
	(0.006)	(0.095)	(0.109)	(0.075)
Sub-estab. distance × $D_{2004-05}$	0.019*	−0.340**	0.163	0.177
	(0.008)	(0.124)	(0.159)	(0.109)
Sub-estab. distance × $D_{2006-10}$	0.033***	−0.399**	0.086	0.313*
	(0.009)	(0.130)	(0.174)	(0.127)
HQ distance × $D_{2000-03}$	−0.009	0.108	−0.059	−0.049
	(0.005)	(0.092)	(0.105)	(0.073)
HQ distance × $D_{2004-05}$	−0.031***	0.285*	−0.161	−0.124
	(0.007)	(0.122)	(0.154)	(0.104)
HQ distance × $D_{2006-10}$	−0.050***	0.316*	−0.111	−0.205^{+}
	(0.008)	(0.129)	(0.169)	(0.121)
R-squared	0.883	0.733	0.817	0.881
Obs.	177,637	177,637	177,637	177,637

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Year_t + \beta_{2,t} \times \text{avg. subordinate establishment distance to MDF}_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the average subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001. The sample only contains subordinate establishments in non-manufacturing firms defined by the HQ's sector classification.

Table B.12: Subordinate establishments by inital skill composition: non-manufacturing

	With high-skilled empl. in 1999			Without high-skilled empl. in 1999		
		Skillshares			Skillshares	
	Employment (1)	Medium (2)	High (3)	Employment (4)	Low (5)	Medium (6)
Sub-estab. distance × $D_{2000-03}$	0.038**	−0.154	0.221	−0.012*	−0.079	−0.127
	(0.013)	(0.250)	(0.254)	(0.006)	(0.115)	(0.127)
Sub-estab. distance × $D_{2004-05}$	0.048*	−0.344	0.454	0.014	−0.525**	0.152
	(0.024)	(0.382)	(0.373)	(0.009)	(0.186)	(0.202)
Sub-estab. distance × $D_{2006-10}$	0.073**	−0.244	0.419	0.018+	−0.499*	−0.029
	(0.027)	(0.416)	(0.400)	(0.011)	(0.194)	(0.224)
HQ distance × $D_{2000-03}$	−0.053***	0.436+	−0.479+	−0.010+	0.076	−0.014
	(0.013)	(0.250)	(0.255)	(0.006)	(0.110)	(0.121)
HQ distance × $D_{2004-05}$	−0.081***	0.720+	−0.785*	−0.019*	0.466*	−0.311
	(0.024)	(0.379)	(0.374)	(0.009)	(0.183)	(0.197)
HQ distance × $D_{2006-10}$	−0.112***	0.550	−0.667+	−0.028**	0.385*	−0.143
	(0.027)	(0.413)	(0.398)	(0.010)	(0.194)	(0.219)
R-squared	0.881	0.877	0.894	0.875	0.739	0.711
Obs.	35,689	35,689	35,689	87,454	87,454	87,454

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree. Samples are split by zero or positive employment of high-skilled labor in 1999. The sample only contains subordinate establishments in non-manufacturing firms defined by the HQ's sector classification.

B.2.6 Sample: Manufacturing

Table B.13: Subordinate establishments: manufacturing

		Skill shares		
	Employment	Low	Medium	High
Sub-estab. distance × $D_{2000-03}$	−0.004	−0.102	0.428*	−0.326*
	(0.011)	(0.163)	(0.213)	(0.148)
Sub-estab. distance × $D_{2004-05}$	−0.024	−0.048	0.253	−0.205
	(0.015)	(0.232)	(0.320)	(0.212)
Sub-estab. distance × $D_{2006-10}$	−0.034+	−0.257	0.245	0.012
	(0.018)	(0.242)	(0.357)	(0.274)
HQ distance × $D_{2000-03}$	−0.004	0.021	−0.386+	0.365*
	(0.011)	(0.161)	(0.209)	(0.147)
HQ distance × $D_{2004-05}$	0.006	−0.098	−0.199	0.297
	(0.015)	(0.226)	(0.315)	(0.213)
HQ distance × $D_{2006-10}$	0.019	0.056	−0.129	0.073
	(0.018)	(0.238)	(0.353)	(0.272)
R-squared	0.942	0.864	0.876	0.899
Obs.	39,309	39,309	39,309	39,309

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Year_t + \beta_{2,t} \times \text{avg. subordinate establishment distance to MDF}_j \times Year_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Year_t$ describes the year dummies from 2000 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the average subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The sample only contains subordinate establishments in manufacturing firms defined by the HQ's sector classification.

Table B.14: Subordinate establishments by inital skill composition: manufacturing

	With high-skilled empl. in 1999			Without high-skilled empl. in 1999		
		Skillshares			Skillshares	
	Employment (1)	Medium (2)	High (3)	Employment (4)	Low (5)	Medium (6)
Sub-estab. distance × $D_{2000-03}$	0.004	0.423	−0.414	0.001	−0.291	0.299
	(0.015)	(0.352)	(0.325)	(0.012)	(0.255)	(0.267)
Sub-estab. distance × $D_{2004-05}$	0.021	−0.146	0.189	-0.038^{+}	−0.099	0.304
	(0.023)	(0.500)	(0.470)	(0.020)	(0.427)	(0.493)
Sub-estab. distance × $D_{2006-10}$	0.026	−0.308	0.250	-0.059^{*}	−0.473	0.820
	(0.033)	(0.529)	(0.486)	(0.025)	(0.416)	(0.535)
HQ distance × $D_{2000-03}$	−0.016	−0.248	0.294	−0.004	0.208	−0.376
	(0.015)	(0.356)	(0.326)	(0.012)	(0.246)	(0.258)
HQ distance × $D_{2004-05}$	-0.048^{*}	0.388	−0.278	0.028	−0.023	−0.448
	(0.024)	(0.503)	(0.473)	(0.020)	(0.411)	(0.484)
HQ distance × $D_{2006-10}$	-0.056^{+}	0.673	−0.389	0.052^{*}	0.246	-0.930^{+}
	(0.033)	(0.533)	(0.491)	(0.025)	(0.401)	(0.526)
R-squared	0.946	0.888	0.904	0.900	0.857	0.817
Obs.	15,280	15,280	15,280	15,819	15,819	15,819

This table shows the results from the fixed-effects difference-in-differences estimation using 1999 as reference year. The estimation equation for the subordinate establishment is: $y_{jt} = \beta_{1,t} \times \text{log(HQ distance to MDF)}_i \times Period_t + \beta_{2,t} \times \text{log(subordinate establishment distance to MDF)}_j \times Period_t + \alpha_j + \alpha_{ct} + \varepsilon_{i(j)t}$, where $Period_t$ describes the time periods from 2000 to 2003, 2004 to 2005, and 2006 to 2010. $\text{log(HQ distance to MDF)}_i$ and $\text{log(subordinate establishment distance to MDF)}_j$ describe the distance of the HQ or the subordinate establishment to their respective MDFs. The table reports the results for $\beta_{1,t}$ and $\beta_{2,t}$. Number of observations reported for each regression. Standard errors clustered at the county-year-level in parentheses. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001. Information on skills reported in social security data. Low-skilled employees do not have any vocational training. Medium-skilled employees have vocational training. High-skilled employees have at least a college degree. Samples are split by zero or positive employment of high-skilled labor in 1999. The sample only contains subordinate establishments in manufacturing firms defined by the HQ's sector classification.

Appendix C

Appendix to Chapter 3

C.1 Data

C.1.1 Data sources and record linkage procedure

Social Security records

Employee history. The Integrated Employment Biographies (*Integrierte Erwerbsbiografien, IEB*) are based on records from the German Social Security System. They contain information on all employees subject to social insurance contributions since 1975 and are updated at least annually. The data cover nearly all private sector employees in Germany, but do not cover civil servants and the self-employed. The IEB contain information on birth year, gender, nationality, education, occupation, full time or part-time status and daily earnings of each employee. Jacobebbinghaus and Seth (2007) and Antoni et al. (2016) provide a detailed description of the structure of the data.[1]

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

Establishment history panel. The Establishment History Panel (*Betriebshistorikpanel, BHP*) is a panel data set that contains information on the sector, number of

[1]The paper by Antoni et al. (2016) focuses on the Sample of Integrated Labor Market Biographies (SIAB), a 2% random sample drawn from the IEB.

employees 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.[2] Sectors are defined based on the Classification of Economic Activities of the German Statistical Office (see also section C.1.2). The location of establishments is provided at the county level. Germany is divided into 402 counties with around 200,000 inhabitants on average. German counties are roughly comparable to counties in the US. Schmucker et al. (2016) provide a detailed description of the data set.

Extension file entry and exit. The extension file entry and exit uses information on worker flows to identify establishment openings and closings. Establishment identifiers may change when a firm restructures. The extension file helps mitigate bias related to restructurings. Hethey and Schmieder (2010) provide details on the file.

ORBIS

We use a linkage table between the Social Security Records and the firm-level database Orbis of the commercial data provider Bureau van Dijk (BvD). BvD compiles its firm-level data from publicly available sources as well as by acquiring data from other commercial data providers. For Germany, BvDs main data provider is Creditreform. Internationally, BvD offers more than 20 different databases with the main customers being privates companies looking for business intelligence on, for instance, competitors, business partners or potential targets for acquisitions. Within BvDs databases, a company is defined 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 at that time. In the years before these changes, a higher share of financial information is missing.

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

Record linkage procedure

The record linkage between Orbis and the Social Security data was performed independently of our project by the German Record Linkage Center (GRLC, see Antoni and Schnell (2017) or www.record-linkage.de for more details on the GRLC). The basis of the linkage was an extract of Orbis acquired by the Institute for Employment Research (IAB). This extract contained data on all German firms at the reference date of January 30, 2014. Of the 1,938,990 firms contained in the data, 1,627,668 were marked as active in Germany at that reference date.

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

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

As these identifiers are non-unique and error-prone, the GRLC developed extensive cleaning, standardization and parsing routines (usually referred to as pre-processing) to achieve records that could successfully be compared between the two data sources.

To deal with remaining differences in, for instance, the spelling or abbreviations of the identifiers, the GRLC applied error-tolerant methods of record linkage (see Christen, 2012). The resulting linkage process consists of 17 consecutive steps, not counting the pre-processing, that varied in terms of which identifiers were used and how strict the requirements on agreement of the compared records were. Schild (2016) provides a more detailed description of the record linkage process. Antoni et al. (2018) report on the linkage success and the representativeness of the resulting data set.

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

Identification of headquarters

The linkage process explained in the previous subsection aimed at identifying as many establishments per firm as possible without trying to determine which of the linked establishments had been the headquarters of the firm. This information was added by the Research Data Centre (FDZ) at the IAB afterwards. To do so, the FDZ performed several iterative steps that mainly relied on the address of the firm according to Orbis and of the establishments according to the administrative data. During later steps the FDZ also used information on the share of administrative staff or the industry code of the establishments under consideration. Given that the administrative data do not contain information that directly identifies an establishment as the headquarters of a superordinate company, this process had to rely on variables that allow the identification of the most likely headquarters among the linked establishments. Antoni et al. (2018) provide more details on the whole process and on the remaining uncertainty regarding the identified headquarters.

C.1.2 Sector and occupation classification

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

The information on the occupation of employees follows the German classification of occupations *"Klassifikation der Berufe"* (KldB). The years 1998-2010 contain the three digit occupation according to the 1988 version of the KldB. The years 2012-2014 contain 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. We therefore exclude 2011 from our analysis.

C.1.3 Assignment of occupations to layers/categories

Layers. To assign occupations to layers, we build on the classification of Caliendo et al. (2015b) for the French PCS ESE occupation classification. We transfer the classification to the international ISCO classification of occupations and from there to the German occupation classification KldB (see section C.1.2). We use official correspondence tables from the German Federal Employment Agency and the International Labor Organization (ILO). In some cases, the translation assigns several layers to the same occupation. Following Friedrich (2016), we generally assign the minimum level of layers to these occupations. Table C.1 displays our assignment of occupations to layers.

Managerial occupations according to Blossfeld (1983, 1987). The assignment treats the following occupations as managerial: 751, 752, 753, 761, 762, 763.

Table C.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
3	751	63124, 71104, 73294, 84394, 94494	Manager, executive, director, board member
2	721, 722, 724, 752, 753, 761, 763, 843	All sub-groups of type 2 in occupation groups: 434, 524, 815; of type 3 in occupation groups: 411, 431, 434, 524, 922; of type 4 in occupation groups: 115, 411, 412, 431, 432, 433, 434, 511, 513, 516, 524, 532, 621, 625, 632, 633, 634, 712, 713, 715, 722, 723, 731, 732, 815, 824, 921, 922, 933; plus: 11494, 21194, 23294, 27194, 27294, 27394, 29194, 29294, 31174, 31194, 41203, 41303, 41383, 41304, 41384, 41394, 41403, 41404, 41484, 41494, 42124, 42144, 42314, 42324, 42394, 43152, 43323, 43343, 43353, 43383, 51133, 51233, 51533, 51543, 51594, 53184, 53394, 61194, 61294, 61394, 63114, 63194, 63313, 71224, 71333, 71433, 72144, 72194, 72243, 73394, 81214, 81234, 81404, 81414, 81424, 81434, 81444, 81454, 81464, 81474, 81484, 81804, 81814, 81884, 82594, 83193, 83194, 84194, 84294, 84304, 84494, 91344, 91354, 92113, 92304, 92394, 92424, 92434, 93303, 93313, 93323, 93343, 93383, 94214, 94493, 94404, 94414, 94484, 94534, 94794	Manager in business organization and strategy, finanical analyst, software developer, qualified IT-specialist, lawyers

Continued on next page

Table C.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
1	31, 32, 601, 602, 603, 605, 606, 607, 611, 612, 621, 622, 623, 625, 626, 627, 628, 629, 762, 811, 813, 841, 842, 844, 862, 863, 871, 872, 873, 874, 875, 881, 882, 883, 604, 624, 633, 687, 812, 822, 831, 851, 852, 853, 855, 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,	Quality manager, training supervisor, management assistant, scientist, engineer, interpreter

Continued on next page

Table C.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples
		33223, 33233, 33243, 33303, 33323, 34203, 34213, 34233, 34303, 34323, 34343, 41213, 41283, 41293, 41322, 41313, 41323, 41314, 41324, 41413, 41423, 41433, 41483, 41414, 41424, 41434, 42114, 42134, 42202, 42334, 43102, 43112, 43122, 43313, 43333, 43363, 51182, 51113, 51123, 51183, 51223, 51243, 51503, 51513, 51523, 51583, 51593, 51504, 51534, 51623, 51663, 53152, 53124, 53134, 53222, 53232, 53312, 53322, 53332, 53314, 61132, 61124, 61204, 61214, 61284, 61314, 62183, 63122, 63132, 63123, 63212, 63213, 71403, 71423, 71522, 71523, 72124, 72134, 72184, 72213, 72223, 72233, 73162, 73163, 73183, 73241, 73202, 73212, 73232, 73242, 73282, 73203, 73213, 73233, 73243, 73253, 73283, 73314, 73324, 73334, 81224, 81294, 81302,81332, 81352, 81382, 81313, 81323, 81333, 81353, 81383, 81393, 81494, 81894, 82212, 82232, 82332, 82343, 82522, 82503, 82523, 82504, 82514, 82524, 82534, 83112, 83132, 83123, 83133, 83124, 83134, 83154, 83223, 84114, 84124, 84134, 84144, 84184, 84214, 84224, 84413, 84404, 84414, 84424, 84434, 84444, 84454, 84484, 91314, 91324, 91334, 91384, 92133, 92384, 92414, 92494, 93213, 93223, 93233, 93333, 93413, 93433, 93513, 93523, 93603, 93613, 93623, 93633, 93643, 93653, 93683, 94224, 94303, 94313, 94323, 94403, 94413, 94483, 94522, 94532, 94582, 94514, 94704, 94714, 94724	
0	Others	Others	Unskilled/semi-skilled occupations in metal-working, printing, machine and equipment assemblers, green keepers, catering, office clerks

Continued on next page

Table C.1: Assignment of occupations to layers

Level	KldB 1988	KldB 2010	Examples

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).

C.1.4 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:

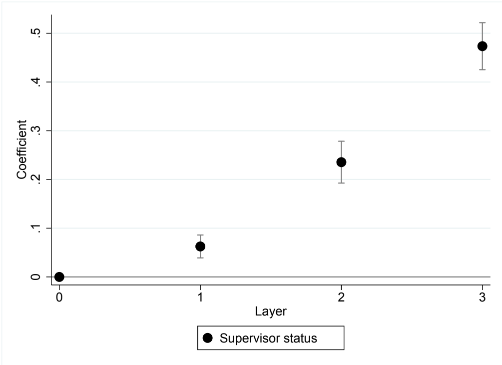
$$y_i = \beta \mathbf{D}_{\text{layer},i} + \gamma \mathbf{X}_i + \delta \mathbf{Z}_i + u_i \tag{C.1}$$

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.

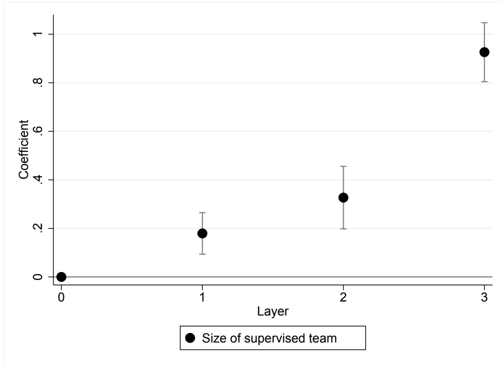
Figure C.1 plots the coefficients and confidence bands by layer. Employees at higher layers are significantly more likely to be supervisors. The predicted probability that an employee at layer 3 is a supervisor at the mean is 84%. Employees at higher layers also supervise larger teams. They are more likely to independently organize their own work. Their duties comprise organizing work for others, making decisions and solving problems. The job of employees at higher layers also require more specific skills. Overall, this descriptive evidence is consistent with the assumption that the assignment of occupations to layers reflects differences between the managerial tasks and duties of employees in firms. Table C.2 presents the estimated coefficients.

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

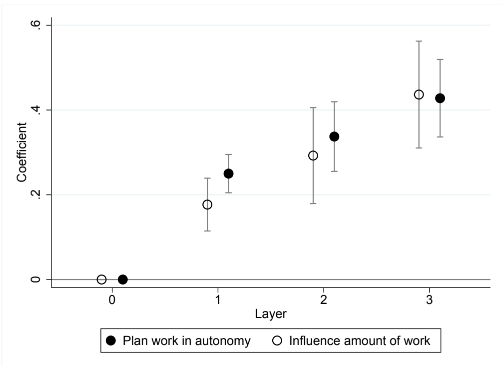
(a) Supervisor status



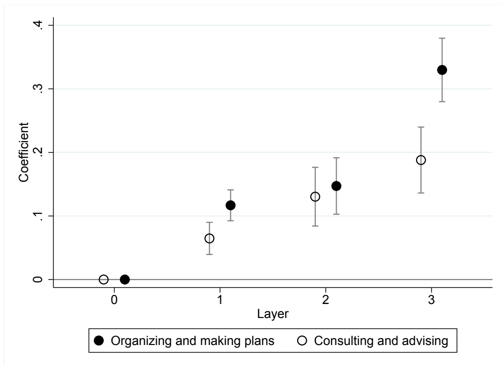
(b) Team size



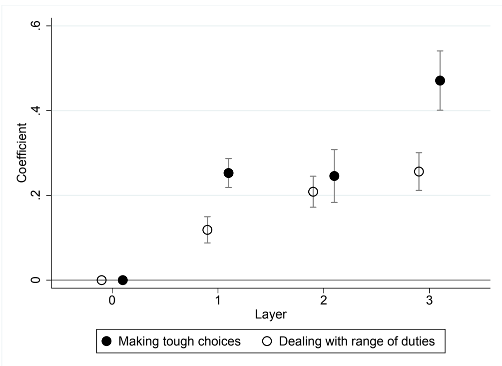
(c) Freely organizing own work



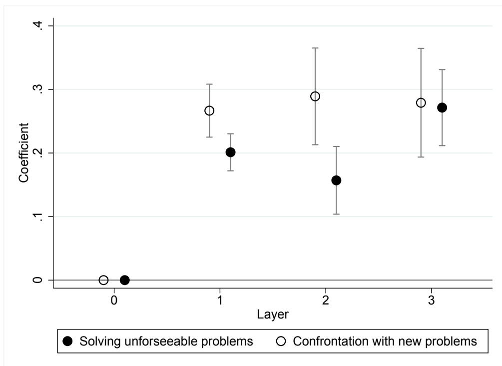
(d) Organizing others' work



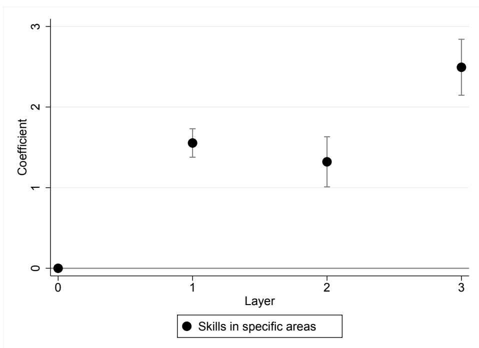
(e) Decision making



(f) Facing and solving problems



(g) Skills



The figure plots the estimated coefficients of the layer dummies in equation (C.1) for different survey questions. See Table C.2 for the survey questions.

Table C.2: Regression results: tasks by layer, 2006 BiBB/BAuA survey

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

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

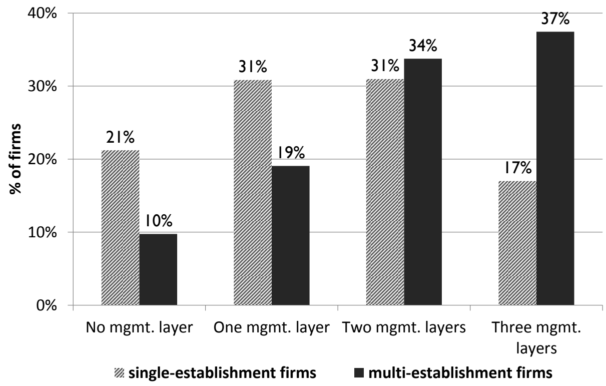
C.2 Facts firm location and organization

Table C.3: Share of firms with consecutive layers, Figure 3.1

# management layers	0	1	2	3
Consecutive organization	Level 0	Level 0+1	Level 0+1+2	Level 0+1+2+3
SE firms	97%	70%	77%	100%
ME firms	91%	56%	72%	100%
Number of firms	21,288	22,115	25,124	18,872

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

Figure C.2: Number of management layers (firm level) by firm type, 2012 cross-section.



The figure plots the distribution of the number of layers separately for SE and ME firms in the 2012 cross-section. The lowest level layer at the firm level is non-managerial. The sample includes firms with at least ten employees.

Table C.4: Regression results, managerial organization (firm level), 2012 cross-section

	# mgmt. layers, Poisson				Manag. share Layers
	(1)	(2)	(3)	(4)	(5)
$D_{\text{ME firm}}$	0.144***	0.111***	0.061***	0.063***	4.247***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.240)
Log # non-mg. employees	0.143***	−0.050***		−0.005	
	(0.002)	(0.011)		(0.003)	
Log # non-mg. employees^2		0.029***			
		(0.002)			
Log sales			0.179***	0.182***	
			(0.002)	(0.003)	
HQ sector FE	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y
# firms	105,948	105,948	53,566	53,566	105,947

2012 cross-section, only firms with at least 10 employees. Robust standard errors in *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, defined at firm level, 5 managerial share in wage sum, layer definition at firm level. *Independent variables:* see Table 3.4. Constant included.

Table C.5: Regression results, managerial organization of ME firms (firm level), 2012 cross-section

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

2012 cross-section, only multi-establishment firms with at least 10 employees. Columns 2, 4, 6 include only ME firms with at least two subordinate establishments. Robust standard errors in parentheses. ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, defined at firm level, 5-6 managerial share in wage sum, firm level layer definition. *Independent variables:* see Table 3.5.

Table C.6: Regression results, managerial organization, 1998-2010 panel

Dependent variable	# mgmt. layers, Poisson				Managerial share, OLS	
					Layers	Blossfeld
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{\text{ME firm}}$	0.045***	0.051***	0.047***	0.010*	1.104***	−0.106+
	(0.002)	(0.002)	(0.004)	(0.004)	(0.121)	(0.060)
Log # non-mg. employees	0.270***	0.341***		0.074***		
	(0.001)	(0.005)		(0.003)		
Log # non-mg. employees^2		−0.008***				
		(0.001)				
Log sales			0.180***	0.144***		
			(0.001)	(0.002)		
HQ sector FE	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y
Legal form FE	N	N	N	N	N	N
# observations	754,578	754,578	101,858	101,858	318,209	318,209

1998-2010 panel, only firms with at least 10 employees in all years. Robust standard errors in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, 5 managerial share in wage sum, layer definition, 6 managerial share in wage sum, Blossfeld. *Independent variables:* see Table 3.4. Constant included. Legal form dummies omitted because of missing information before 2005.

Table C.7: Regression results, managerial organization of ME firms, 1998-2010 panel

Dependent variable	# mgmt. layers, Poisson				Managerial share, OLS			
					Layers		Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maximum log	0.053***		0.038***		1.916*		0.303***	
distance to HQ	(0.001)		(0.003)		(0.044)		(0.021)	
Log area		0.022***		0.016***		0.612***		0.076**
spanned by firm		(0.001)		(0.002)		(0.030)		(0.015)
Log # non-mg.	0.196***	0.170***						
employees	(0.001)	(0.002)						
Log sales			0.111***	0.087***				
			(0.002)	(0.003)				
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y
# firms	85,899	30,710	15,277	6,978	85,899	30,710	85,899	30,710

1998-2010 panel, only ME firms with at least 10 employees in all years. Columns 2, 4, 6, 8 include only ME firms with at least two subordinate establishments. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-4 number of management layers, 5-6 managerial share in wage sum, layer definition, 7-8 managerial share in wage sum, Blossfeld. *Independent variables:* see Table 3.5. Constant included. Legal form dummies omitted because of missing information before 2005.

Table C.8: Regression results, managerial organization, 2012 cross-section, by legal form

# mgmt. layers, Poisson	GmbH & Co. KG			GmbH		
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{\text{ME firm}}$	0.082***	0.029	0.016	0.088***	0.018*	0.024**
	(0.015)	(0.021)	(0.021)	(0.006)	(0.008)	(0.008)
Log # non-mg. employees	0.224***		0.028*	0.141***		−0.011**
	(0.005)		(0.009)	(0.002)		(0.003)
Log sales		0.220***	0.206***		0.188***	0.193***
		(0.006)	(0.008)		(0.002)	(0.003)
HQ sector dummies	Y	Y	Y	Y	Y	Y
HQ county dummies	Y	Y	Y	Y	Y	Y
# firms	18,653	9,242	9,242	84,203	42,468	42,468

# mgmt. layers, Poisson	AG		
	(7)	(8)	(9)
$D_{\text{ME firm}}$	0.037*	0.049**	0.048*
	(0.017)	(0.019)	(0.020)
Log # non-mg. employees	0.057***		0.001
	(0.005)		(0.007)
Log sales		0.065***	0.065***
		(0.004)	(0.006)
HQ sector dummies	Y	Y	Y
HQ county dummies	Y	Y	Y
# firms	2,823	1,635	1,635

2012 cross-section, only firms with at least 10 employees, by legal form. A "GmbH & Co. KG" is a limited partnership with a limited liability company as general partner. A "GmbH" is a limited liability company. An "AG" is a public company. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* number of management layers. *Independent variables:* see Table 3.4.

Table C.9: Regression results, managerial organization, 2012 cross-section, robustness

Dependent variable	# mgmt. layers, Poisson			Managerial share, OLS					
					Layers			Blossfeld	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$D_{\text{ME firm}}$	0.078***	0.085***	0.089***	1.669***	2.218***	2.060***	1.838***	1.300***	1.273***
	(0.008)	(0.006)	(0.006)	(0.306)	(0.268)	(0.235)	(0.180)	(0.142)	(0.124)
Log # non-mg.	0.152***	0.148***	0.153***						
employees	(0.002)	(0.002)	(0.002)						
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
HQ county FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Legal form FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# firms	101,161	103,695	105,543	101,159	103,693	105,541	101,159	101,159	105,541

2012 cross-section, only multi-establishment firms with at least 10 employees. Columns 1, 4, 7 use only ME firms with all establishments in the same sector. Columns 2, 5, 8 use only ME firms where all subordinate establishments are located outside the headquarters county. Columns 3, 6, 9 use only ME firms with size smaller than the 95th percentile of the SE firm size distribution. Robust standard errors in parentheses. *** $p < 0.001$. *Dependent variable:* 1-3 number of management layers, 4-6 managerial share in wage sum, layer definition, 7-9 managerial share in wage sum, Blossfeld. *Independent variables:* see Table 3.4.

Table C.10: Regression results, # management layers in ME firms, 2012 cross-section

Dependent variable	# mgmt. layers Poisson		Manag. share, layers, OLS		Manag. share, Blossfeld, OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{\text{2. quartile of max. log distance}}$	0.014		-1.222^{+}			-0.384
	(0.015)		(0.701)			(0.313)
$D_{\text{3. quartile of max. log distance}}$	0.068^{***}		1.186			0.516
	(0.015)		(0.791)			(0.355)
$D_{\text{4. quartile of max. log distance}}$	0.093^{***}		3.113^{***}			1.555^{***}
	(0.015)		(0.866)			(0.391)
$D_{\text{2. quartile of log area}}$		0.048^{*}		-1.199		-0.183
		(0.024)		(1.264)		(0.602)
$D_{\text{3. quartile of log area}}$		$0.118{*}{*}{*}$		0.884		1.420^{*}
		(0.025)		(1.474)		(0.695)
$D_{\text{4. quartile of log area}}$		0.138^{***}		3.658^{*}		2.627^{***}
		(0.025)		(1.659)		(0.797)
Log # non-managerial employees	0.138^{***}	0.116^{***}				
	(0.004)	(0.006)				
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
# firm-years	9,275	3,320	9,275	3,320	9,275	3,320

2012 cross-section, only multi-establishment firms with at least 10 employees. Columns 2, 4, 6 include only ME firms with at least two subordinate establishments. Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* 1-2 number of management layers, 3-4 managerial share in wage sum, layer definition, 5-6 managerial share in wage sum, Blossfeld. *Independent variables:* $D_{\text{X. quartile of max. log distance}}$: Xth quartile of log of maximum distance between subordinate establishment and headquarters; $D_{\text{X. quartile of log area}}$: Xth quartile of log of minimum area covered by establishments in square kilometers; others see Table 3.4.

C.3 A model of multi-establishment firm organization

Assumption 1. *The predictability of the production process* λ*, the communication costs* θ_{j0} *and the learning costs* c *are such that*

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

C.3.1 The optimal organization of a single establishment firm

Lagrangian equation and first order conditions

We use equation (3.9), which is binding in optimum, to substitute for $n^{\ell}_{0,L}$, $\ell > 0$:

$$\begin{aligned}
\mathcal{L} &= n^0_{0,L} w_0 \left(1 + c z^0_{0,L}\right) + n^0_{0,L} \sum_{\ell=1}^{L-1} \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} w_0 \left(1 + c z^{\ell}_{0,L}\right) + w_0 \left(1 + c \bar{z}_{0,L}\right) \\
&\quad + \xi_{0,L} \left(\tilde{q} - n^0_{0,L} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right)\right) + \varphi_{0,L} \left(n^0_{0,L} \theta_{00} e^{-\lambda z^{L-1}_{0,L}} - 1\right) \\
&\quad + \bar{\eta}^L_{0,L} (z^{L-1}_{0,L} - \bar{z}_{0,L}) + \sum_{\ell=1}^{L-1} \bar{\eta}^{\ell}_{0,L} (z^{\ell-1}_{0,L} - z^{\ell}_{0,L}) - \bar{\eta}^0_{0,L} z^0_{0,L} - \eta^0_{0,L} n^0_{0,L} \\
\frac{\partial \mathcal{L}}{\partial \bar{z}_{0,L}} &= w_0 c - \xi_{0,L} n^0_{0,L} \lambda e^{-\lambda \bar{z}_{0,L}} - \bar{\eta}^L_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial z^{L-1}_{0,L}} &\begin{cases} \overset{L=1}{=} n^0_{0,1} \left(w_0 c - \varphi_{0,1} \theta_{00} \lambda e^{-\lambda z^0_{0,1}}\right) + \bar{\eta}^1_{0,1} - \bar{\eta}^0_{0,1} = 0 \\ \overset{L\geq 1}{=} n^0_{0,L} \left(w_0 c \theta_{00} e^{-\lambda z^{L-2}_{0,L}} - \varphi_{0,L} \theta_{00} \lambda e^{-\lambda z^{L-1}_{0,L}}\right) + \bar{\eta}^L_{0,L} - \bar{\eta}^{L-1}_{0,L} = 0 \end{cases} \\
\frac{\partial \mathcal{L}}{\partial z^{\ell}_{0,L}} &= n^0_{0,L} w_0 \left(c \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} - \lambda \theta_{00} e^{-\lambda z^{\ell}_{0,L}} (1 + c z^{\ell+1}_{0,L})\right) - \bar{\eta}^{\ell}_{0,L} + \bar{\eta}^{\ell+1}_{0,L} = 0 \\
&\text{for } 0 < \ell < L-1,\ L > 2 \\
\frac{\partial \mathcal{L}}{\partial z^0_{0,L}} &\overset{L\geq 1}{=} n^0_{0,L} w_0 \left(c - \lambda \theta_{00} e^{-\lambda z^0_{0,L}} (1 + c z^1_{0,L})\right) + \bar{\eta}^1_{0,L} - \bar{\eta}^0_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial n^0_{0,L}} &= w_0 \left[\left(1 + c z^0_{0,L}\right) + \sum_{\ell=1}^{L-1} \theta_{00} e^{-\lambda z^{\ell-1}_{0,L}} w_0 \left(1 + c z^{\ell}_{0,L}\right)\right] \\
&\quad - \xi_{0,L} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right) + \varphi_{0,L} \theta_{00} e^{-\lambda z^{L-1}_{0,L}} - \eta^0_{0,L} = 0 \\
\frac{\partial \mathcal{L}}{\partial \xi_{0,L}} &= \tilde{q} - n^0_{0,L} \left(1 - e^{-\lambda \bar{z}_{0,L}}\right) = 0
\end{aligned}$$

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

Proposition 1: Comparative statics

We focus on $L \in \{1,2,3\}$ because these are relevant for the empirics. The second order conditions are given by:

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

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

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

for $0 < \ell < L-1$, $L > 2$

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

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

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

$$\frac{\partial^2\mathcal{L}}{\partial\varphi_{0,L}\partial\tilde{q}} = \frac{dn^0_{0,L}}{d\tilde{q}}\theta_{00}e^{-\lambda z^{L-1}_{0,L}} - n^0_{0,L}\theta_{00}\lambda e^{-\lambda z^{L-1}_{0,L}}\frac{dz^{L-1}_{0,L}}{d\tilde{q}} = 0$$

where we substitute $\frac{d\mathcal{L}}{dz^\ell_{0,L}}$, $\ell < L$, into equation $\frac{d^2\mathcal{L}}{dn^L_{0,L}d\tilde{q}}$.

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

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

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 $\frac{d^2\mathcal{L}}{dn_{0,L}^0 d\tilde{q}}$:

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

4. From $\frac{d^2\mathcal{L}}{dz_{0,L}^{L-1}d\tilde{q}}$, with $\frac{d^2\mathcal{L}}{dz_{0,L}^{\ell}d\tilde{q}}$, $\ell < L-1$:

$$\begin{aligned}
\frac{d\varphi_{0,1}}{d\tilde{q}} &= \varphi_{0,1}\lambda\frac{dz_{0,1}^0}{d\tilde{q}} \equiv \varphi_{0,1}\lambda f_1\left(\varphi_{0,L}\right)\frac{dz_{0,L}^{L-1}}{d\tilde{q}}\\
\frac{d\varphi_{0,2}}{d\tilde{q}} &= \varphi_{0,2}\lambda\frac{dz_{0,2}^1}{d\tilde{q}}(1-\theta_{00}e^{-\lambda z_{0,2}^0}) \equiv \varphi_{0,2}\lambda f_2\left(\varphi_{0,L}\right)\frac{dz_{0,L}^{L-1}}{d\tilde{q}}\\
\frac{d\varphi_{0,3}}{d\tilde{q}} &= \varphi_{0,3}\lambda\frac{dz_{0,3}^2}{d\tilde{q}}\left(1-\frac{\frac{dz_{0,3}^1}{d\tilde{q}}}{\frac{dz_{0,3}^2}{d\tilde{q}}}\right) \equiv \varphi_{0,3}\lambda f_3\left(\varphi_{0,L}\right)\frac{dz_{0,L}^{L-1}}{d\tilde{q}}
\end{aligned}$$

with $f_L\left(\varphi_{0,L}\right) > 0$ for $L = 1, 2, 3$.

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

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

To show (a): The number $n_{0,L}^{\ell}$ and the knowledge $z_{0,L}^{\ell}$ of employees at all below-CEO layers $\ell < L$ increase with total 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. $\ell = L-1$, $L > 1$: $\frac{dn_{0,L}^{L-1}}{d\tilde{q}} = \frac{dn_{0,L}^0}{d\tilde{q}}\theta_{00}e^{-\lambda z_{0,L}^{L-2}}f_L(\varphi_{0,L}) > 0$ by $f_L(\varphi_{0,L}) > 0$.
3. $L-1 > \ell > 0$, $L > 2$: analogously to $\ell = L-1$.

Knowledge of employees:

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

To show (b): The marginal benefit of CEO time $\varphi_{0,L}$ and the marginal production cost $\xi_{0,L}$ increase with total output $\tilde{q}$.

$\frac{d\varphi_{0,L}}{d\tilde{q}} > 0$ follows from $\frac{d\varphi_{0,L}}{d\tilde{q}} = \varphi_{0,L}\lambda f_L(\varphi_{0,L})\frac{dz_{0,L}^{L-1}}{d\tilde{q}} > 0$ by $f_L(\varphi_{0,L}) > 0$ and $\frac{dz_{0,L}^{L-1}}{d\tilde{q}} > 0$. Substituting into $\frac{d\xi_{0,L}}{d\tilde{q}}$ yields:

$$\frac{d\xi_{0,L}}{d\tilde{q}} > 0 \quad \text{if} \quad \varphi_{0,L}f_L(\varphi_{0,L})\theta_{00}e^{-\lambda z_{0,L}^{L-1}} > \xi_{0,L}e^{-\lambda \bar{z}_{0,L}} = \frac{w_0 c}{\lambda}\theta_{00}e^{-\lambda z_{0,L}^{L-1}}.$$

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

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

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

C.3.2 The optimal organization of a multi-establishment firm

Lagrangian equation and first order conditions, single output market

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

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

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

Establishment-level: The number and knowledge of employees We use equation (3.21), which is binding in optimum, to substitute for $n^{\ell}_{j,L}$, $\ell > 0$. L_j denotes the number of below-CEO layers at establishment j.

$$\begin{aligned}
\mathcal{L} = {} & n^0_{j,\omega} w_j \left(1 + c z^0_{j,\omega}\right) + n^0_{j,\omega} \sum_{\ell=1}^{L_j} \theta_{00} e^{-\lambda z^{\ell-1}_{j,\omega}} w_j \left(1 + c z^{\ell}_{j,\omega}\right) + s_{j,\omega} w_0 \left(1 + c\bar{z}_{0,\omega}\right) \\
& + \xi_{j,\omega}\left(q_{j,\omega} - n^0_{j,\omega}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right)\right) + \varphi_{j,\omega}\left(n^0_{j,\omega}\theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - s_{j,\omega}\right) \\
& + \bar{\eta}^L_{j,\omega}(z^{L_j}_{j,\omega} - \bar{z}_{0,\omega}) + \sum_{\ell=1}^{L_j} \bar{\eta}^{\ell}_{j,\omega}(z^{\ell-1}_{j,\omega} - z^{\ell}_{j,\omega}) - \bar{\eta}^0_{j,\omega} z^0_{j,\omega} - \eta^0_{j,\omega} n^0_{j,\omega}
\end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial z^{L_j}_{j,\omega}} \begin{cases} \overset{L_j=0}{=} n^0_{j,\omega}\left(w_j c - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^0_{j,\omega}}\right) + \bar{\eta}^L_{j,\omega} - \bar{\eta}^0_{j,\omega} = 0 \\ \overset{L_j>0}{=} n^0_{j,\omega}\left(w_j c\theta_{00} e^{-\lambda z^{L_j-1}_{j,\omega}} - \varphi_{j,\omega}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}}\right) + \bar{\eta}^L_{j,\omega} - \bar{\eta}^{L_j}_{j,\omega} = 0 \end{cases}$$

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

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

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

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial n^0_{j,\omega}} = {} & w_j\left[\left(1 + c z^0_{j,\omega}\right) + \sum_{\ell=1}^{L_j} \theta_{00} e^{-\lambda z^{\ell-1}_{j,\omega}} w_j \left(1 + c z^{\ell}_{j,\omega}\right)\right] \\
& - \xi_{j,\omega}\left(1 - e^{-\lambda \bar{z}_{0,\omega}}\right) + \varphi_{j,\omega}\theta_{j0} e^{-\lambda z^{L_j}_{j,\omega}} - \eta^0_{j,\omega} = 0
\end{aligned}$$

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

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

Endogenous variables:

$$
\begin{aligned}
e^{\lambda z_{j,\omega}^{L_j}} &= \frac{q_{j,\omega}}{1-e^{-\lambda \bar{z}_{0,\omega}}} \frac{\theta_{j0}}{s_{j,\omega}} \\
e^{\lambda\left(z_{j,\omega}^{\ell-1}-z_{j,\omega}^{\ell-2}\right)} &= \left(1+cz_{j,\omega}^{\ell}\right)\lambda \quad \forall \ell = 2,...,L_j, \\
e^{\lambda z_{j,\omega}^{0}} &= \left(1+cz_{j,\omega}^{1}\right)\lambda\theta_{jj} \\
\xi_{j,\omega} &= \frac{w_j\left(1+cz_{j,\omega}^{0}+\frac{1}{\lambda}+\mathbb{1}(L_j\geq 1)\frac{\theta_{00}}{\lambda}\sum_{\ell=1}^{L_j} e^{-\lambda z_{j,\omega}^{\ell-1}}\right)}{1-e^{-\lambda \bar{z}_{0,\omega}}}, \\
\varphi_{j,\omega} &= \frac{w_j c}{\lambda\theta_{j0}}\theta_{00} e^{\lambda\left(z_{j,\omega}^{L_j}-z_{j,\omega}^{L_j-1}\right)} \quad \text{for } L_j>0, \quad \varphi_{j,\omega} = \frac{w_j c}{\lambda\theta_{j0}} e^{\lambda z_{j,\omega}^{0}} \quad \text{for } L_j = 0.
\end{aligned}
$$

Proposition 2: Allocation of output and CEO time

Proof. The first order conditions imply:

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

$$
\begin{aligned}
\phi_{0,\omega} = \frac{\partial C}{\partial q_{0,\omega}} - \bar{\phi}_{0,\omega} &= \frac{\partial C}{\partial q_{1,\omega}} - \bar{\phi}_{0,\omega} = \phi_{1,\omega} = 0 \quad \text{for } q_0, q_1 > 0 \\
\Rightarrow \quad \frac{\partial C}{\partial q_{0,\omega}} &= \xi_{0,\omega} = \xi_{1,\omega} = \frac{\partial C}{\partial q_{1,\omega}}.
\end{aligned}
$$

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

$$
\begin{aligned}
\kappa_{0,\omega} &= \frac{\partial C}{\partial s_{0,\omega}} - w_0(1+c\bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} = \\
\kappa_{1,\omega} &= \frac{\partial C}{\partial s_{1,\omega}} - w_0(1+c\bar{z}_{0,\omega}) + \bar{\kappa}_{0,\omega} = 0 \quad \text{for } s_0, s_1 > 0 \\
\Rightarrow \quad \frac{\partial C}{\partial s_{0,\omega}} &= \varphi_{0,\omega} = \varphi_{1,\omega} = \frac{\partial C}{\partial s_{1,\omega}}.
\end{aligned}
$$

Proof of Corollary 1. Proposition 2 requires that both the marginal costs of production $\xi_{j,\omega}$ and the marginal benefit of CEO time $\varphi_{j,\omega}$ are equal if the firm produces at two establishments. $\varphi_{j,\omega}$ is a function of θ_{10}, but $\xi_{j,\omega}$ is not. Production at two establishments with the same number of below-CEO layers L_j and symmetric communication costs $\theta_{10} = \theta_{00}$ but wages $w_1 \geq w_0$ therefore violates Proposition 2.

To see this, consider the case $L_j = 0\,\forall j$. The following two equations cannot be fulfilled at the same time, where $w_0 = w_1 = w$, because the first requires that the knowledge levels are the same, the second requires that they are different:

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

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

L_j denotes the number of below-CEO layers at establishment j. The second order conditions for $\tilde{q} > q_{j,\omega} > 0\,\forall j$, $1 > s_{j,\omega} > 0\,\forall j$ are given by:

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

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

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

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

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

$$\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega}d\tilde{q}} \begin{cases} \overset{L_j=0}{=} -\frac{d\varphi_{j,\omega}}{d\tilde{q}}\theta_{j0}\lambda e^{-\lambda z^0_{j,\omega}} + \varphi_{j,\omega}\theta_{j0}\lambda^2 e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\tilde{q}} = 0 \\ \overset{L_j>0}{=} -w_j c\theta_{00}\lambda e^{-\lambda z^{L_j-1}_{j,\omega}}\frac{dz^{L_j-1}_{j,\omega}}{d\tilde{q}} - \frac{d\varphi_{j,\omega}}{d\tilde{q}}\theta_{j0}\lambda e^{-\lambda z^{L_j}_{j,\omega}} + \varphi_{j,\omega}\theta_{j0}\lambda^2 e^{-\lambda z^{L_j}_{j,\omega}}\frac{dz^{L_j}_{j,\omega}}{d\tilde{q}} = 0 \end{cases}$$

$$\frac{\partial^2\mathcal{L}}{\partial z^\ell_{j,\omega}\partial\tilde{q}} = -\lambda\theta_{00}ce^{-\lambda z^{\ell-1}_{j,\omega}}\frac{dz^{\ell-1}_{j,\omega}}{d\tilde{q}} + \lambda^2\theta_{00}e^{-\lambda z^\ell_{j,\omega}}\frac{dz^\ell_{j,\omega}}{d\tilde{q}}(1 + cz^{\ell+1}_{j,\omega}) - \lambda\theta_{00}e^{-\lambda z^\ell_{j,\omega}}c\frac{dz^{\ell+1}_{j,\omega}}{d\tilde{q}} = 0$$

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

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\tilde{q}} = \lambda^2\theta_{00}e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\tilde{q}}(1 + cz^1_{j,\omega}) - \lambda\theta_{00}e^{-\lambda z^0_{j,\omega}}c\frac{dz^1_{j,\omega}}{d\tilde{q}} = 0 \quad \text{for } L_j > 0$$

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

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

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

where we substitute $\frac{d\mathcal{L}}{dz^\ell_{j,\omega}}$, $\ell < L$, into equation $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\tilde{q}}$.

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

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

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

1. As will be shown below, $\frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0$ and $\frac{dz_{j,\omega}^{\ell}}{d\tilde{q}} = 0 \, \forall \ell < L$. Thus, $\frac{d^2\mathcal{L}}{dn_{j,\omega}^{0} d\tilde{q}}$ yields:

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

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

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

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

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

To show (a): The knowledge of the employees at all below-CEO layers $z_{j,\omega}^{\ell}$, $\forall \ell < L$ is constant.

We know: $\xi_{0,\omega} = \xi_{1,\omega}$ and $\varphi_{0,\omega} = \varphi_{1,\omega}$. These equations uniquely determine the knowledge levels (two equations in two unknowns). They do not depend on $\tilde{q}$, so the knowledge levels do not depend on $\tilde{q}$.

$\Rightarrow \frac{dz_{j,\omega}^{\ell}}{d\tilde{q}} = 0 \forall \ell, j$

To show (b): The share of CEO time $s_{j,\omega}$ and the number of employees at all below-CEO layers $n_{j,\omega}^{\ell}$ at the location with relatively lower (higher) wages increase (decreases) with total output $\tilde{q}$, where the threshold ratio of wages depends on ω. Local output $q_{j,\omega}$ increases if the share of CEO time $s_{j,\omega}$ does.

We focus on $L \in \{1, 2, 3\}$, because these are relevant for the empirics.

1. $\frac{ds_{j,\omega}}{d\tilde{q}} > 0$ if $f(w_j) < f(w_k)\, k \neq j$. For $L_0 = L_1 = 0$, $\frac{ds_{0,\omega}}{d\tilde{q}} > 0$ if $w_0 < w_1$. For $L_0 = L_1 = 1$, $\frac{ds_{0,\omega}}{d\tilde{q}} > 0$ if $w_0 e^{\lambda(z^0_{1,\omega} - z^0_{0,\omega})} < w_1$. For $L_0 = 0, L_1 = 1$, $\frac{ds_{0,\omega}}{d\tilde{q}} > 0$ if $w_0 \frac{1}{\theta_{00}} e^{\lambda z^0_{1,\omega}} < w_1$. For $L_0 = 1, L_1 = 0$, $\frac{ds_{0,\omega}}{d\tilde{q}} > 0$ if $w_0 \theta_{00} e^{-\lambda z^0_{0,\omega}} < w_1$. Analogous results hold for $L = 3$.
2. $\text{sgn}\left(\frac{dn^\ell_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{dn^0_{j,\omega}}{d\tilde{q}}\right)$ and $\text{sgn}\left(\frac{dn^0_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{ds_{j,\omega}}{d\tilde{q}}\right)$, i.e. the number of employees varies as the share of CEO time.
3. $\text{sgn}\left(\frac{dq_{j,\omega}}{d\tilde{q}}\right) = \text{sgn}\left(\frac{ds_{j,\omega}}{d\tilde{q}}\right)$ for $\frac{ds_{j,\omega}}{d\tilde{q}} > 0$. By $\frac{dq_{0,\omega}}{d\tilde{q}} + \frac{dq_{1,\omega}}{d\tilde{q}} = 1$ and $\frac{ds_{0,\omega}}{d\tilde{q}} + \frac{ds_{1,\omega}}{d\tilde{q}} = 0$, the sign is indeterminate if $\frac{ds_{j,\omega}}{d\tilde{q}} < 0$.

To show (c): The marginal benefit of CEO time $\varphi_{j,\omega}$ does not vary with total output $\tilde{q}$.
From $\frac{d^2\mathcal{L}}{dz^{L_j}_{j,\omega} d\tilde{q}}$:

$$\frac{d\varphi_{j,\omega}}{d\tilde{q}} = \varphi_{j,\omega}\lambda \frac{dz^0_{j,\omega}}{d\tilde{q}} \text{ if } L_j = 0$$
$$\frac{d\varphi_{j,\omega}}{d\tilde{q}} = \varphi_{j,\omega}\lambda \frac{dz^{L_j}_{j,\omega}}{d\tilde{q}} - w_0 c e^{\lambda(z^{L_j}_{j,\omega} - z^{L_j - 1}_{j,\omega})} \frac{dz^{L_j-1}_{j,\omega}}{d\tilde{q}} \text{ if } L_j > 0$$

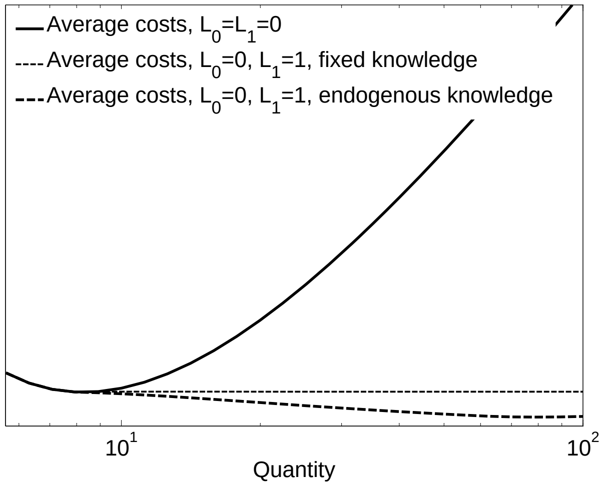
$\Rightarrow \frac{d\varphi_{j,\omega}}{d\tilde{q}} = 0 \forall j$ by $\frac{dz^\ell_{j,\omega}}{d\tilde{q}} = 0 \forall j, \ell$

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

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

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

Figure C.3: Illustration: Proof of Proposition 4.



Parameter values: $\frac{c}{\lambda} = .225$, $\theta_{10} = \theta_{00} = .26$ (from Caliendo and Rossi-Hansberg, 2012), $w_0 = w_1 = 1$.

Proposition 4: The optimal number of layers

a) To show: The average costs of the $\{L_j/L_j\}$-organization coincide with the average costs of a single establishment firm with L layers characterized in Proposition 1c): The average cost function of the $\{L_j/L_j\}$-organization is U-shaped and reaches a minimum at $\tilde{q}_L^*$.

The firm with a $\{L_j/L_j\}$-organization chooses symmetric knowledge levels by $\xi_{0,\omega} = \xi_{1,\omega}$ and $\varphi_{0,\omega} = \varphi_{1,\omega}$. The cost function thus coincides with the cost function of a single-establishment firm given $\sum_{j=0}^{1} n_{j,\omega}^0 = n_{0,L_j+1}^0$. Correspondingly, Proposition 1 applies.

b) To show: The average costs of the $\{L_j/L_j+1\}$-organization are lower than the average costs of the $\{L_j/L_j\}$-organization for output levels $\tilde{q} > \tilde{q}_L^*$.

That is, consider an ME firm with L_j below-CEO layers at both establishments at the minimum efficient scale $\tilde{q}_L^*$. There exists a range of quantities $\tilde{q} > \tilde{q}_L^*$ such that the average cost of an ME firm with L_j below-CEO layers at establishment j and $L_j + 1$ below-CEO layers at establishment $k \neq j$ are lower than the minimum average cost of an ME firm with L_j below-CEO layers at both establishments.

For simplicity and without loss of generality, we choose $j = 0$, $k = 1$ and $L_j = 0$.

We proceed in two steps. Figure C.3 illustrates the argument.

1. We construct an ME organization with $L_0 = 0$ below-CEO layers at establishment 0 and $L_1 = 1$ below-CEO layers at establishment 1 that has the same average cost as an ME organization with $L_1 = L_0 = 0$ below-CEO layers at both establishments at the minimum efficient scale $\tilde{q}_1^*$.

   Consider an ME firm with $L_1 = L_0 = 0$. By $w_1 = w_0, \theta_{10} = \theta_{00}$, its cost function coincides with the cost function of a SE firm with one layer $L = 1$. At the minimum efficient scale $\tilde{q}_1^*$,

$$\xi_{0,1} = AC_{0,1} \equiv AC_{0,1}^{MES} \tag{C.2}$$

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

$$\lambda \bar{z}_{0,1} = \lambda z_{0,1}^0 + \ln\left(\lambda z_{0,1}^0 + \frac{\lambda}{c} + 1 + \theta_{00} e^{-\lambda z_{0,1}^0}\right) - \ln\theta_{00} \tag{C.4}$$

$$\tilde{q}_{0,1}^* = \frac{1}{\theta_{00}} e^{\lambda z_{0,1}^0}(1 - e^{-\lambda \bar{z}_{0,1}}) \tag{C.5}$$

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

$$z_{0,\omega}^0 = z_{0,1}^0 \tag{C.6}$$

$$\bar{z}_{0,\omega} = \bar{z}_{0,1} \tag{C.7}$$

$$w_0\left(1 + cz_{0,1}^0 + \frac{c}{\lambda}\right) = w_1\left(1 + cz_{1,\omega}^0 + \frac{c}{\lambda} + \frac{c}{\lambda}\theta_{11}e^{-\lambda z_{1,\omega}^0}\right), \quad \text{i.e. } \xi_{1,\omega} = \xi_{0,1}, \tag{C.8}$$

$$\text{and } \frac{w_0}{\theta_{00}} e^{\lambda z_{0,1}^0} = \frac{w_1\theta_{11}}{\theta_{10}} e^{\lambda(z_{1,\omega}^1 - z_{1,\omega}^0)}, \quad \text{i.e. } \varphi_{1,\omega} = \varphi_{0,1}, \tag{C.9}$$

$$\text{with } z_{1,\omega}^1 = \frac{1}{\lambda\theta_{11}} e^{\lambda z_{1,\omega}^0} - \frac{1}{c}.$$

   By construction, the average cost of the ME firm at $\tilde{q}_1^*$ are $AC_{0,\omega} = AC_{0,1}^{MES}$. □

2. We show that the average cost of an ME firm with organization $\omega = \{L_0 = 0, L_1 = 1\}$ and optimal knowledge levels are lower than the minimum average costs of the ME organization with L_j layers at both establishments for $\tilde{q} > \tilde{q}_1^*$, because they are lower than the average cost an ME firm with

organization $\omega = \{L_0 = 0, L_1 = 1\}$ and fixed knowledge levels.

The maximum producible quantity $\tilde{q}^{MAX}$ of the ME firm with organization $\omega = \{L_0 = 0, L_1 = 1\}$ and fixed knowledge levels is given by

$$\tilde{q}^{MAX} = \frac{1}{\theta_{10}} e^{\lambda z^1_{1,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}}) \tag{C.10}$$

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

$$\xi_{1,\omega} = \xi_{0,\omega} = \xi_{0,1} \quad \text{by construction} \tag{C.11}$$

$$AC_{0,\omega} = \frac{w_1 \left(1 + c z^0_{1,\omega} + \frac{c}{\lambda}\right) + \theta_{10} e^{-\lambda z^1_{1,\omega}} w_0 (1 + c \bar{z}_{0,\omega})}{1 - e^{-\lambda \bar{z}_{0,\omega}}} \tag{C.12}$$

$$= \xi_{0,\omega} - \frac{\frac{w_1 c}{\lambda} \theta_{11} e^{-\lambda z^0_{1,\omega}} - \theta_{10} e^{-\lambda z^1_{1,\omega}} w_0 (1 + c \bar{z}_{0,\omega})}{1 - e^{-\lambda \bar{z}_{0,\omega}}}$$

$$= \xi_{0,\omega} - w_1 \theta_{00} \frac{c}{\lambda} e^{-\lambda z^0_{1,\omega}} e^{-\lambda z^0_{0,\omega}} \left(e^{\lambda z^0_{0,\omega}} - \theta_{00} \left(\frac{\lambda}{c} + \lambda \bar{z}_{0,\omega} \right) \right) \text{ by } \varphi_{0,\omega} = \varphi_{1,\omega} \tag{C.13}$$

$$= \xi_{0,\omega} = AC^{MES}_{0,1} \quad \text{by (C.3)} \tag{C.14}$$

i.e. the ME firm produces both $\tilde{q}^*_1$ and $\tilde{q}^{MAX}$ at the same average costs.

The ME firm produces quantities $\tilde{q}$ with $\tilde{q}^{MAX} \geq \tilde{q} \geq \tilde{q}^*_1$ by allocating the share s to the establishment with one below-CEO layer and the share $1 - s$ of the production quantity to the establishment with two below-CEO layers, where

$$s = \frac{\tilde{q} - \frac{1}{\theta_{10}} e^{\lambda z^1_{1,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}})}{\frac{1}{\theta_{00}} e^{\lambda z^0_{0,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}}) - \frac{1}{\theta_{10}} e^{\lambda z^1_{1,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}})} \tag{C.15}$$

Both numerator and denominator are negative. The denominator is constant. $0 \leq s \leq 1$, because the numerator achieves its minimum at $\tilde{q} = \frac{1}{\theta_{00}} e^{\lambda z^0_{0,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}})$ (so $s = 1$), and its maximum at $\tilde{q} = \frac{1}{\theta_{10}} e^{\lambda z^1_{1,\omega}} (1 - e^{-\lambda \bar{z}_{0,1}})$ (so $s = 0$). That is, the average cost function of the ME firm with fixed knowledge levels is flat for $\tilde{q} \in \left[\tilde{q}^*_1, \tilde{q}^{MAX}\right]$ (see the light dashed line in Figure C.3).

The average cost of an ME firm with organization $\omega = \{L_0 = 0, L_1 = 1\}$ 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.3) because

$$C(\tilde{q}) \leq C(\tilde{q}, \bar{z}_{0,\omega}, z^0_{0,1}, z^0_{1,\omega}(z^0_{0,1}), z^1_{1,\omega}(z^0_{0,1})) \tag{C.16}$$

Consequently, there exist quantities $\tilde{q} > \tilde{q}^*_1$ such that the average cost of an ME firm with organization $\omega = \{L_0 = 0, L_1 = 1\}$ are lower than the average cost of an ME firm with $L_0 = 0$ below-CEO layers at both establishments, as well as an SE firm with $L = 1$. □

c) To show: The average cost function of the $\{L_j+1/L_j+1\}$-organization intersects the average cost function of the $\{L_j/L_j\}$-organization at the output $\tilde{q}^{L+1}_L$, with $q^*_{L+1} > \tilde{q}^{L+1}_L > q^*_L$. It intersects the average cost function of the $\{L_j/L_j+1\}$-organization at the output $\tilde{q} > \tilde{q}^{L+1}_L$.

That is, the average cost function of an ME organization with L_j below-CEO layers at establishment j and L_j+1 below-CEO layers at establishment k intersect the average cost function of an organization with L_j+1 below-CEO layers at both establishments at a higher quantity than the average cost function of an organization with L_j below-CEO layers at both establishments does.

We exploit the characteristics of the average cost function.

- $AC_{0,\omega} \leq AC^{MES}_{0,1} \, \forall \, \tilde{q}^*_1 \leq \tilde{q} \leq \tilde{q}^{MAX}$,
- $AC_{0,1}$ is increasing for $\tilde{q} > \tilde{q}^*_1$,
- $AC_{0,2}$ is decreasing for $\tilde{q} \leq \tilde{q}^*_2$, where $\tilde{q}^{MAX} \leq \tilde{q}^*_2$,
- at $\tilde{q}^*_1$, $AC_{0,2} > AC_{0,1}$.

In consequence, the increasing average costs function of the ME firm with $L_j = 0$ below-CEO layers $AC_{0,1}$ intersects the decreasing average costs function of the ME firm with $L_j = 1$ below CEO layers at both establishments $AC_{0,2}$ at a lower quantity than the (weakly) decreasing average cost function of the ME firm with organization $\omega = \{L_0 = 0, L_1 = 1\}$ $AC_{0,\omega}$ intersects the average cost function $AC_{0,2}$. □

Proposition 5: Comparative statics with respect to θ_{10}

The second order conditions for $\tilde{q} > q_{j,\omega} > 0\,\forall j$, $1 > s_{j,\omega} > 0\,\forall j$ are given by:

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

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

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

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

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

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

$$\frac{d^2\mathcal{L}}{dz^{L_1}_{1,\omega}d\theta_{10}}\begin{cases}\overset{L_1=0}{=} -\frac{d\varphi_{1,\omega}}{d\theta_{10}}\theta_{10}\lambda e^{-\lambda z^0_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^0_{1,\omega}}\frac{dz^0_{1,\omega}}{d\theta_{10}} - \varphi_{1,\omega}\lambda e^{-\lambda z^0_{1,\omega}} = 0\\ \overset{L_1>0}{=} -w_1 c\lambda\theta_{00}e^{-\lambda z^{L_1-1}_{1,\omega}}\frac{dz^{L_1-1}_{1,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}}\lambda\theta_{10}e^{-\lambda z^{L_1}_{1,\omega}} + \varphi_{1,\omega}\theta_{10}\lambda^2 e^{-\lambda z^{L_1}_{1,\omega}}\frac{dz^{L_1}_{1,\omega}}{d\theta_{10}}\\ -\varphi_{1,\omega}\theta_{10}\lambda e^{-\lambda z^{L_1}_{1,\omega}} = 0\end{cases}$$

$$\frac{\partial^2\mathcal{L}}{\partial z^\ell_{j,\omega}\partial\theta_{10}} = -\lambda\theta_{00}ce^{-\lambda z^{\ell-1}_{j,\omega}}\frac{dz^{\ell-1}_{j,\omega}}{d\theta_{10}} + \lambda^2\theta_{00}e^{-\lambda z^\ell_{j,\omega}}\frac{dz^\ell_{j,\omega}}{d\theta_{10}}(1 + cz^{\ell+1}_{j,\omega}) - \lambda\theta_{00}e^{-\lambda z^\ell_{j,\omega}}c\frac{dz^{\ell+1}_{j,\omega}}{d\theta_{10}} = 0$$

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

$$\frac{d^2\mathcal{L}}{dz^0_{j,\omega}d\theta_{10}} = \lambda^2\theta_{00}e^{-\lambda z^0_{j,\omega}}\frac{dz^0_{j,\omega}}{d\theta_{10}}(1 + cz^1_{j,\omega}) - \lambda\theta_{00}e^{-\lambda z^0_{j,\omega}}c\frac{dz^1_{j,\omega}}{d\theta_{10}} = 0 \quad \text{for } L_j > 0$$

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

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

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

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

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

where we substitute $\frac{d\mathcal{L}}{dz^{\ell}_{j,\omega}}$, $\ell < L$, into equation $\frac{d^2\mathcal{L}}{dn^0_{j,\omega}d\theta_{10}}$.

To show (a): The total number of employees at all below-CEO layers $\sum_{j=0}^{1} n^{\ell}_{j,\omega}$, $\forall \ell < L$ decreases with the communication costs θ_{10}.

- $\ell = 0$: Follows from $\frac{d^2\mathcal{L}}{d\xi_{j,\omega}d\theta_{10}}$, with $-\frac{dq_{0,\omega}}{d\theta_{10}} - \frac{dq_{1,\omega}}{d\theta_{10}} = 0$:

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

- $\ell > 0$: 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 (a): The knowledge of the CEO $\bar{z}_{0,\omega}$ increases with the communication costs θ_{10}.

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

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

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

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

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

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

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

$$\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} = \frac{\varphi_{0,\omega}\theta_{00}}{\lambda\xi_{0,\omega}(1+e^{-\lambda \bar{z}_{0,\omega}})(\theta_{00} e^{\lambda z^{L_1}_{1,\omega}} - \theta_{10} e^{\lambda z^{L_0}_{0,\omega}})} > 0 \quad \text{for sufficiently low } w_1$$

To show (a): The knowledge of the employees at all below-CEO layers $z^{\ell}_{j,\omega}$, $\forall \ell < L$ increases with the communication costs θ_{10}.

Follows from $\frac{d^2\mathcal{L}}{dz_{j,\omega}^{L_j} d\theta_{10}}$ and $\frac{\partial^2\mathcal{L}}{\partial z_{j,\omega}^{\ell}\partial\theta_{10}}$, $\frac{\partial^2\mathcal{L}}{\partial z_{j,\omega}^{0}\partial\theta_{10}}$ by $\frac{d\varphi_{0,\omega}}{d\theta_{10}} > 0$.

To show (b): The marginal benefit of CEO time $\varphi_{j,\omega}$ increases with the communication costs θ_{10}.

$\frac{d^2\mathcal{L}}{dn_{0,\omega}^0 d\tilde{q}}$ yields, together with $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$ and $\frac{d\bar{z}_{0,\omega}}{d\theta_{10}} > 0$:

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

To show (b): The marginal production cost $\xi_{j,\omega}$ increases with the communication costs θ_{10}.

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

To show: The increase of the below-CEO knowledge levels with the communication costs θ_{10} is stronger at higher than at lower layers and at the subordinate establishment than at the headquarters.

- Higher vs. lower layers: From $\frac{d^2\mathcal{L}}{dz_{j,\omega}^{L_j} d\theta_{10}}$, $L_j > 0$:

$$\frac{dz_{j,\omega}^{L_j}}{d\theta_{10}} - \frac{dz_{j,\omega}^{L_j-1}}{d\theta_{10}} = \frac{d\varphi_{j,\omega}}{d\theta_{10}} \frac{1}{\lambda \varphi_{j,\omega}} + \mathbf{1}(j=1)\varphi_{1,\omega}\theta_{10}\lambda e^{-\lambda z_{1,\omega}^{L_1}} > 0$$

- Subordinate establishment vs. headquarters: From $\frac{d^2\mathcal{L}}{dz_{0,\omega}^{L_0} d\theta_{10}}$ and $\frac{d^2\mathcal{L}}{dz_{1,\omega}^{L_1} d\theta_{10}}$ with $\frac{d\varphi_{0,\omega}}{d\theta_{10}} - \frac{d\varphi_{1,\omega}}{d\theta_{10}} = 0$ and $\varphi_{0,\omega} = \varphi_{1,\omega}$:

$$\frac{dz_{1,\omega}^0}{d\theta_{10}} = \frac{dz_{0,\omega}^0}{d\theta_{10}} + \frac{1}{\lambda\theta_{10}}$$

Proposition 6: Allocation of production quantity and CEO time (extension)

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

$$- \mathbf{1}(q_{1,\omega} \geq \tilde{q}_1 \wedge q_{0,\omega} \leq \tilde{q}_0)\underline{\phi}_{0,\omega}(q_{1,\omega} - \tilde{q}_1 + \tau(q_{0,\omega} - \tilde{q}_0))$$

First-order conditions:

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

Implications:

$$\begin{aligned}
\phi_{j,\omega} = 0 \forall j \quad &\Rightarrow \quad \xi_{1,\omega} = \tau\xi_{0,\omega} \text{ if } q_{0,\omega} > \tilde{q}_0 \wedge q_{1,\omega} < \tilde{q}_1 \\
&\qquad \xi_{0,\omega} = \tau\xi_{1,\omega} \text{ if } q_{1,\omega} > \tilde{q}_1 \wedge q_{0,\omega} < \tilde{q}_0 \\
&\qquad \frac{1}{\tau}\xi_{1,\omega} < \xi_{0,\omega} < \tau\xi_{1,\omega} \text{ if } q_{0,\omega} = \tilde{q}_0 \wedge q_{1,\omega} = \tilde{q}_1 \\
\exists \phi_{j,\omega} > 0 \quad &\Rightarrow \quad \xi_{j,\omega} > \tau\xi_{-j,\omega} \text{ at } \tilde{q}_{j,\omega} = 0
\end{aligned}$$

C.3.3 Proposition 7: Optimal output

Using the implicit function theorem, we show:

$$\begin{aligned}
\frac{d\tilde{q}(\alpha_i)}{d\alpha_i} &= \frac{(R_0 + R_1)\left(\frac{\sigma}{\sigma-1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma}}{1 + \alpha_i\left(R_0 + R_1\right)\left(\frac{\sigma}{\sigma-1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma-1} \frac{\sigma^2}{\sigma-1}\frac{d\xi_{0,\omega}}{d\tilde{q}}} \qquad > 0 \\
\frac{d\tilde{q}(\alpha_i)}{d\theta_{10}} &= -\frac{\alpha_i\left(R_0 + R_1\right)\left(\frac{\sigma}{\sigma-1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma-1} \frac{\sigma^2}{\sigma-1}\frac{d\xi_{0,\omega}}{d\theta_{10}}}{1 + \alpha_i\left(R_0 + R_1\right)\left(\frac{\sigma}{\sigma-1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma-1} \frac{\sigma^2}{\sigma-1}\frac{d\xi_{0,\omega}}{d\tilde{q}}} \qquad < 0 \quad \text{by } \frac{d\xi_{0,\omega}}{d\theta_{10}} > 0
\end{aligned}$$

The denominator is positive by

$$\begin{aligned}
&1 + \alpha_i\left(R_0 + R_1\right)\left(\frac{\sigma}{\sigma-1}\xi_{0,\omega}\left(\tilde{q}(\alpha_i)\right)\right)^{-\sigma-1}\frac{\sigma^2}{\sigma-1}\frac{d\xi_{0,\omega}}{d\tilde{q}} \\
=&1 - \frac{\tilde{q}}{\frac{\sigma}{\sigma-1}\xi_{0,\omega}}\frac{\sigma^2}{\sigma-1}\frac{\xi_{0,\omega}\lambda e^{-\lambda\bar{z}_{0,\omega}}}{1 - e^{-\lambda\bar{z}_{0,\omega}}}\frac{1 - e^{-\lambda\bar{z}_{0,\omega}}}{\lambda\tilde{q}(1 + e^{-\lambda\bar{z}_{0,\omega}})}
\end{aligned}$$

$$=1 - \frac{\sigma e^{-\lambda \bar{z}_{0,\omega}}}{1+e^{-\lambda \bar{z}_{0,\omega}}} \quad > 0 \quad \text{for } \sigma < e^{\lambda \bar{z}_{0,\omega}} + 1.$$

C.4 Organizational response to new high speed train routes

Table C.11: Reduction of travel times in minutes through high speed routes

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

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

Table C.12: Robustness check, managerial share, treated establishments

Establishment managerial share	(1)	(2)	(3)	(4)
$D_{\theta\downarrow,ijt}$	-2.856^{*}	-0.643^{*}	-0.614	-0.669^{*}
	(1.086)	(0.364)	(0.387)	(0.288)
R-squared	0.872	0.855	0.867	0.855
Establishment FE	Y	Y	Y	Y
County-year FE	Y	Y	Y	Y
Observations	2,399	37,798	201,895	37,798

2000-2010 panel, only firms with at least 10 employees. Standard errors clustered at county level in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Dependent variable:* share of employees of establishment in managerial occupations. *Independent variable:* see Table 3.8. Column 1: Only years 2000, 2004, and 2008. Column 2: control group includes establishments with travel time reduction of less than 30 minutes. Column 3: control group of column 2 plus establishments in counties without ICE train station. Column 4: $D_{\theta\downarrow,ijt}$ defined as equal to 1 if travel time to headquarters is reduced by at least 10 minutes and zero otherwise.

Bibliography

Acemoglu, D., P. Aghion, R. Griffith, and F. Zilibotti (2010): "Vertical Integration and Technology: Theory and Evidence," *Journal of the European Economic Association*, 8, 989–1033.

Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 4, 1043–1171.

Acemoglu, D. and P. Restrepo (2017): "Robots and Jobs: Evidence from US Labor Markets," *NBER Working Paper No.w23285.*

Akerman, A., I. Gaarder, and M. Mogstad (2015): "The Skill Complementarity of Broadband Internet," *The Quarterly Journal of Economics*, 130, 1781–1824.

Antoni, M., A. Ganzer, and P. vom Berge (2016): "Sample of Integrated Labour Market Biographies (SIAB) 1975-2014," FDZ-Datenreport 04/2016.

Antoni, M., K. Koller, M.-C. Laible, and F. Zimmermann (2018): "Orbis-ADIAB: From Record Linkage Key to Research Dataset. Combining Commercial Company Data with Administrative Employer-Employee Data," FDZ-Methodenreport, forthcoming.

Antoni, M. and R. Schnell (2017): "The Past, Present and Future of the German Record Linkage Center (GRLC)," *Jahrbücher für Nationalökonomie und Statistik*, forthcoming.

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

Atasoy, H. (2013): "The Effects of Broadband Internet Expansion on Labor Market Outcomes," *Industrial and Labor Relations Review*, 66, 315–345.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

Baker, G. P. and T. N. Hubbard (2003): "Make vs. Buy in Trucking: Asset Ownership, Job Design, and Information," *The American Economic Review*, 93, 551–572.

——— (2004): "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," *The Quarterly Journal of Economics*, 119, 1443–1479.

Baldwin, R. (2016): *The Great Convergence: Information Technology and the New Globalization*, The Belknap Press of Harvard University Press.

Battiston, D., J. Blanes i Vidal, and T. Kirchmaier (2017): "Face-to-Face Communication in Organisations," *CEPR Discussion Paper No. 11924.*

Bernard, A. and J. B. Jensen (2007): "Firm Structure, Multinationals, and Manufacturing Plant Deaths," *Review of Economics and Statistics*, 89, 193–204.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2017): "Production Networks, Geography and Firm Performance," *Journal of Political Economy*, forthcoming.

Bertschek, I., W. Briglauer, K. Hüschelrath, B. Kauf, and T. Niebel (2015): "The Economic Impacts of Broadband Internet: A Survey," *Review of Network Economics*, 14, 201–227.

Bertschek, I., D. Cerquera, and G. J. Klein (2013): "More Bite - More Bucks? Measuring the Impact of Broadband Internet on Firm Performance," *Information Economics and Policy*, 25, 190–203.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2017): "What Drives Differences in Management?" *NBER Working Paper No. 23300.*

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying (2015): "Does Working from Home Work? Evidence from a Chinese Experiment," *The Quarterly Journal of Economics*, 130, 165–218.

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

——— (1987): "Labor-Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany," *American Journal of Sociology*, 93, 89–118.

Brux, G. (2006): "NBS Nürnberg - Ingolstadt in Betrieb," *Der Eisenbahningenieur*, 142–143.

Brynjolfsson, E. and L. M. Hitt (2000): "Beyond Computation: Information Technology, Organizational Transformation and Business Performance," *The Journal of Economic Perspectives*, 14, 23–48.

Brynjolfsson, E. and A. McAffee (2011): *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*, Digital Frontier Press.

——— (2014): *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company.

Brynjolfsson, E. and A. Saunders (2010): *Wired for Innovation: How Information Technology Is Reshaping the Economy*, The MIT Press, Cambridge, Massachusetts.

Bundesagentur für Arbeit (2016): "Arbeitsmarkt in Zahlen: Beschäftigungsstatistik," Version of 2016.

Bundesnetzagentur (2011): "Annual Report," Tech. rep.

——— (2017): "Borders of Local Service Areas 2017," Web.

Cairncross, F. (1997): *The Death of Distance: How the Communications Revolution Will Change Our Lives*, Harvard Business School Press.

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, 809–852.

Caliendo, L. and E. Rossi-Hansberg (2012): "The Impact of Trade on Organization and Productivity," *The Quarterly Journal of Economics*, 127, 1393–1467.

Canzian, G., S. Poy, and S. Schueller (2015): "Broadband Diffusion and Firm Performance in Rural Areas: Quasi-Experimental Evidence," *IZA Discussion Paper No. 9429.*

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Cardona, M., T. Kretschmer, and T. Strobel (2013): "ICT and Productivity: Conclusions from the Empirical Literature," *Information Economics and Policy*, 25, 109–125.

Chandler, A. D. (2002): *The Visible Hand: The Managerial Revolution in American Business*, Cambridge, MA & London: Harvard University Press.

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

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

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

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

Columbia Telecommunications Corporation (2010): "The Impact of Broadband Speed and Price on Small Business," Tech. rep., Small Business Administration, Office of Advocacy.

Crèmer, J., L. Garicano, and A. Prat (2007): "Language and the Theory of the Firm," *The Quarterly Journal of Economics*, 122, 373–407.

Czernich, N., O. Falck, T. Kretschmer, and L. Woessmann (2011): "Broadband Infrastructure and Economic Growth," *The Economic Journal*, 121, 505–532.

De Stefano, T., R. Kneller, and J. Timmis (2014): "The (Fuzzy) Digital Divide: The Effect of Broadband Internet Use on UK Firm Performance," *University of Nottingham Discussion Papers in Economics No. 14/06.*

Deutscher Bundestag (2007): "Flughafenanbindungen von Fernverkehrshaltepunkten der Bahn," Tech. rep., BT-Drucksache 16/7104.

Draca, M., R. Sadun, and J. Van Reenen (2006): "Productivity and ICT: A Review of the Evidence," *Centre for Economic Performance, London School of Economics and Political Science, London, UK.*

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 124, 843–881.

Eberle, J., P. Jacobebbinghaus, J. Ludsteck, and J. Witter (2011): "Generation of Time-Consistent Industry Codes in the Face of Classification Changes: Simple Heuristic Based on the Establishment History Panel (BHP)," FDZ-Methodenreport 05/2011.

Eurailpress.de (2002): "NBS Köln-Rhein/Main wird heute eröffnet," 24 January 2018.

——— (2004): "Hamburg - Berlin: Hochgeschwindigkeitsstrecke feierlich eröffnet," 24 January 2018.

——— (2006): "Premierenfahrt auf der Schnellbahnachse Nürnberg-Ingolstadt-München," 24 January 2018.

——— (2007): "Schnelle Bahnverbindung: Lufthansa streicht Flüge Köln - Frankfurt," 24 January 2018.

Falck, O., R. Gold, and S. Heblich (2014): "E-lections: Voting Behavior and the Internet," *The American Economic Review*, 104, 2238–2265.

Federal Ministry of Economics and Technology (2009): "Broadband Atlas Germany," Tech. rep.

Federal Ministry of Transport and Digital Infrastructure (2017): "Broadband Deployment in Germany," 5 October 2017.

Fitzenberger, B., A. Osikominu, and R. Völter (2005): "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample," *Schmollers Jahrbuch*, 126, 405–436.

Fort, T. C. (2017): "Technology and Production Fragmentation: Domestic Versus Foreign Sourcing," *The Review of Economic Studies*, 84, 650–687.

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, 874–904.

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

Gaspar, J. and E. Glaeser (1998): "Information Technology and the Future of Cities," *Journal of Urban Economics*, 43, 136–156.

Giroud, X. (2013): "Proximity and Investment: Evidence from Plant-level Data," *The Quarterly Journal of Economics*, 128, 861–915.

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, 105–127.

Gumpert, A. (2018): "The Organization of Knowledge in Multinational Firms," *Journal of the European Economic Association*, forthcoming.

Gumpert, A., H. Steimer, and M. Antoni (2018): "Firm Organization with Multiple Establishments," *mimeo, LMU Munich.*

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," Tech. rep., GESIS Cologne, Germany (data access); Federal Institute of Vocational Education and Training, Bonn doi:10.4232/1.4820.

Heining, J., W. Klosterhuber, P. Lehnert, and S. Seth (2016): "Linked Employer-Employee Data from the IAB: LIAB Longitudinal Model 1993-2014 (LIAB LM 9314)," *FDZ*, 10/2016.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): "Export versus FDI with Heterogeneous Firms," *The American Economic Review*, 94, 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.

Iacus, S., G. King, and G. Porro (2009): "CEM: Software for Coarsened Exact Matching," *Journal of Statistical Software*, 30, 1–27.

Jacobebbinghaus, P. and S. Seth (2007): "The German Integrated Employment Biographies Sample IEBS," *Schmollers Jahrbuch*, 127, 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, 157–179.

Kandilov, I. T. and M. Renkow (2010): "Infrastructure Investment and Rural Economic Development: An Evaluation of USDA's Broadband Loan Program," *Growth and Change*, 41, 165–191.

Koehler, M. (2012): "Internet-Verbindungen als Standortfaktor erster Güte," *Frankfurter Allgemeine Zeitung*, on August 29.

Kolko, J. (2012): "Broadband and Local Growth," *Journal of Urban Economics*, 71, 100–113.

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

Nueno, P. and P. Ghemawat (2002): "Revitalizing Philips," Ivey Publishing.

OECD (2005): "Communications Outlook 2005," Tech. rep.

OpenStreetMap (2017): "Planet Dump Retrieved from https://planet.osm.org," Web.

QGIS Development Team (2017): *QGIS Geographic Information System*, Open Source Geospatial Foundation.

Rajan, R. G. and J. Wulf (2006): "The Flattening Firm: Evidence From Panel Data on the Changing Nature of Corporate Hierarchies," *The Review of Economics and Statistics*, 88, 759–773.

Reses, J. (2013): "Memorandum to Yahoo Employees," 22 February 2013.

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

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

Schnabel, P. (2015): *Kommunikationstechnik-Fibel*, vol. 4, Elektronik-Kompendium.de.

Stockinger, B. (2017): "The Effect of Broadband Internet on Establishments' Employment Growth: Evidence from Germany," *IAB Discussion Paper No. 201719*.

Storper, M. and A. J. Venables (2004): "Buzz: Face-to-Face Contact and the Urban Economy," *Journal of Economic Geography*, 4, 351–371.

Zuckerberg, M. (2016): "Mark Zuckerberg Live Video at Facebook HQ (Introducing New Facebook Office Inside)," 1 January 2016.