Rising Labour Market Inequality: Regional Disparities and Wage-Setting Institutions

Terry Gregory

University of Regensburg

Rising Labour Market Inequality: Regional Disparities and Wage-Setting Institutions

DISSERTATION

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von:
Dipl.-Vw. Terry Gregory

Berichterstatter:
Prof. Dr. Dr. h.c. Joachim Möller
JProf. Dr. Melanie Arntz

Tag der Disputation: 26.1.2015

Acknowledgment

After finishing this long thesis, it is a pleasure to devote some lines to all the people that I owe so many thanks. First of all, I would like to thank Joachim Möller for supervising my PhD thesis. I am very grateful for receiving so many helpful and constructive comments during various doctoral seminars in Regensburg. I always enjoyed the atmosphere in the seminars and left Regensburg with a long list of valuable suggestions on how to improve my work. Many thanks also to all his PhD students and postdocs that stimulated an exciting discussion during these events. Thank you also to Jürger Jerger for beeing an excellent chairman during my defence in Regensburg.

A special warm thank you goes to Melanie Arntz who did excellent on-site supervision at ZEW. I really enjoyed the spirit of some Friday evening discussions when ZEW turned into a ghost house while we discussed scientific problems for hours, drawing wild pictures on white boards, and generating a continuous series of aha!-effects. This surely taught me how to think through and eventually solve an economic problem. I guess the fact that we share the same "Rheinish" mentality also helped provide us with an optimum communication base. I really feel lucky to have such a great supervisor and mentor.

I also benefited from an excellent research environment at ZEW. My thanks go to the entire department for numerous discussions, helpful feedback and an inspiring atmosphere. Special thanks to Stephan Dlugosz, Bernd Fitzenberger and Francois Laisney for providing helpful assistance whenever simple OLS regressions did not do the job. I also owe many thanks to Jan Melcher, Rahel Felder and Helge Matthiessen for excellent student assistance. My gratitude also goes to Holger Bonin and the ZEW management for all their support and trust during my PhD. We often take things for granted, but we shouldn't.

Furthermore, I would like to thank my co-authors Roberto Patuelli, Florian Lehmer and Bodo Aretz for their contributions that helped to shape our joint research. Moreover, I would like to thank all listeners of several conferences and workshops that provided helpful comments and suggestions on my research. Also, thanks to Ulrich Zierahn, Anna Salomons, Konrad Stahl, Michael Meier and Sebastian Butschek for reading parts of my thesis and coming up with helpful suggestions on how to improve my work.

Of course, money is not everything, but without money, all is nothing. My research would not have been possible without the financial support of the Fritz-Thyssen Foundation, German Research Foundation, ZEW Sponsors' Association for Science and Practice, Ministry of Labour and Social Affairs and several ZEW internal funding opportunities for which I feel very grateful. Note that the authors are responsible for all results and conclusions derived in this dissertation and do not necessarily reflect the views of the sponsors.

I would also like to thank my family and friends, especially my parents. My mother's discipline and my father's optimistic American way of life has always been something that I took as an example and which shaped my thought probably more than I would admit – thank you so much! Finally, I owe a big thank you to my beloved wife who always supported me with personnel advice and understanding in times when work demanded extra hours. Her patience and support during the end of my thesis, when I was finalizing several papers at the same time, was beyond what one can hope for in challenging times. The least I can do is dedicate this thesis to my biggest supporter.

Mannheim, August 2014

Terry Gregory

To Luisa

Contents

List of Tables

List of Figures

Introduction

This cumulative dissertation comprises five research articles that pursue two distinct research agendas on the topic of labour market inequality. The first part deals with the role of geography for inequality and particularly stresses the role of demographic ageing and geographical migration in reinforcing regional disparities. The second part deals with wage-setting institutions, where the focus lies on the economic effects of minimum wages (MWs) on employment and earnings (inequality). The two topics are motivated in the following.

Part I: Demographic Change and Regional Labour Market Disparities

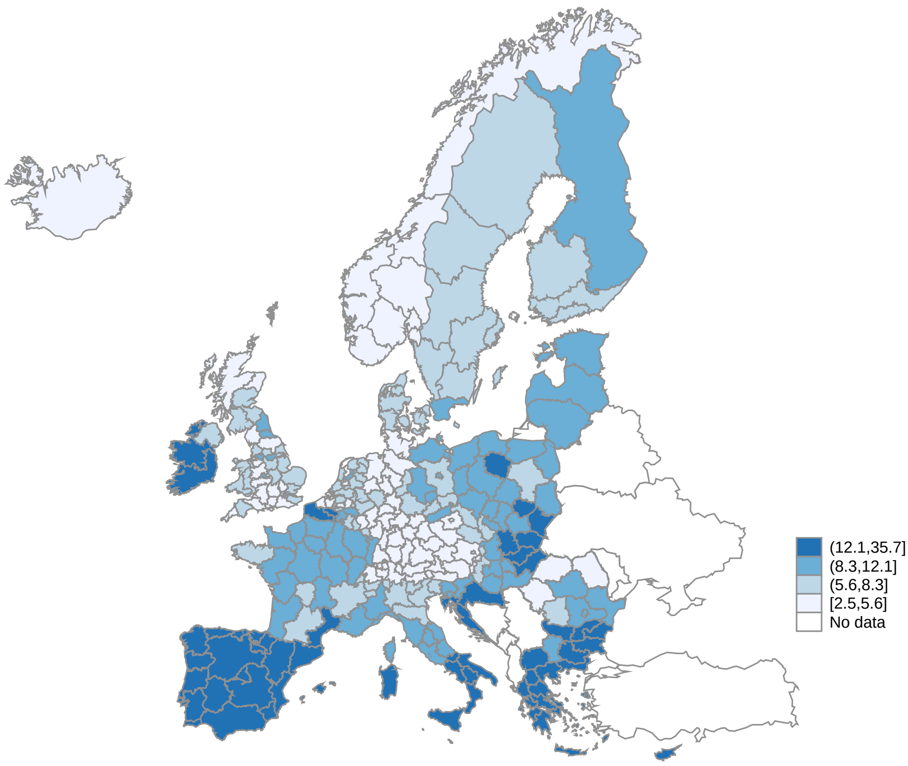
One of the most remarkable trends among advanced economies is the large and growing disparities between regional labour markets. For the US, Moretti (2012) documents a "great divergence" between communities that he describes as one of the most fundamental changes in the economy: While few regions with a well-educated labour force and a strong innovation sector are increasingly successful in creating new jobs and offering high wages, other regions with a less-educated workforce and depressed industries increasingly fall behind. As the main reason for this divide, the author sees long-run economic forces that are changing the economy in profound ways and which have a lot to do with the transition to a knowledge-based economy. More than traditional industries, the knowledge economy tends to be geographically concentrated, implying that initial regional conditions matter considerably for future development. Regions that are already successful tend to further attract young and educated workers, thus triggering a cumulative process towards more polarized regions. As a result, the salary of workers increasingly depends on where you live rather than on your personal characteristics. More generally, inequality in advanced economies to a large extent reflects a geographical divide.

Such divergence processes can also be observed for other advanced economies such as Germany. While few cities in Eastern Germany such as Berlin, Jena and Dresden have been improving economically and are increasingly able to attract young and highly qualified labour, other areas are suffering from low economic performance and out-migration of their young and talented workforce, leaving them depopulated. The latter indicates that Eastern Germany's catching-up was to a large extent a geographically uneven process. In addition, a similar polarization trend can also be observed between West German regions, as the thesis shows. Moreover, increasing regional disparities may be relevant beyond the German case. In fact, Figure 1 maps the most recent unemployment rates for European NUTS-2 regions. For instance, the first quintile (light blue) depicts the 20% regions with the highest unemployment rate in 2013, whose values range from 2.5 to 5.6%. The fifth quintile (dark blue) contains the 20% regions with the lowest corresponding value. The map shows that whereas the unemployment rate is around 2.5% in some European countries, others reveal unemployment rates of up to 35.7%. These large disparities might seem surprising given several studies which show a convergence between European member states. However, despite this convergence, differences between regions *within* European states have actually been increasing (OECD, 2005). In fact, these geographic inequalities pose a key challenge for EU regional and territorial cohesion policy (European Parliament, 2007). Of course, the reasons for these disparities and divergence processes are complex and partly country-specific. However, we will discuss some important determinants throughout this thesis for the case of German regions that should be relevant in a more European context as well.

Understanding the role of agglomeration forces and knowledge spillovers in exacerbating existing disparities is an important issue in this context. A lot has to do with the fact that most highly educated workers live in innovation hubs where they earn high salaries and have "good" jobs, while low educated workers are left in regions with backward oriented industries. However, wages are thereby higher in some regions than others due to the sorting of high educated workers into these regions. Although this is interesting, it is less surprising. More important, high educated workers not only earn higher salaries, but also have a favourable impact on their surrounding regions. The reasons are local multipliers arising, for instance, through knowledge spillovers within and between firms in the local economy (for a recent overview, see Moretti 2011). Since such multipliers tend to be most relevant in creative industries, not all cities profit from such externalities though.

The observation of diverging regions may seem somewhat surprising given the influential work by O'brien (1992) Caimcross (1997) and Friedman (2005). According to their view, location

Figure 1: Unemployment rates for European NUTS-2 regions (2013)



Notes: Own illustrations based on data from Eurostat.

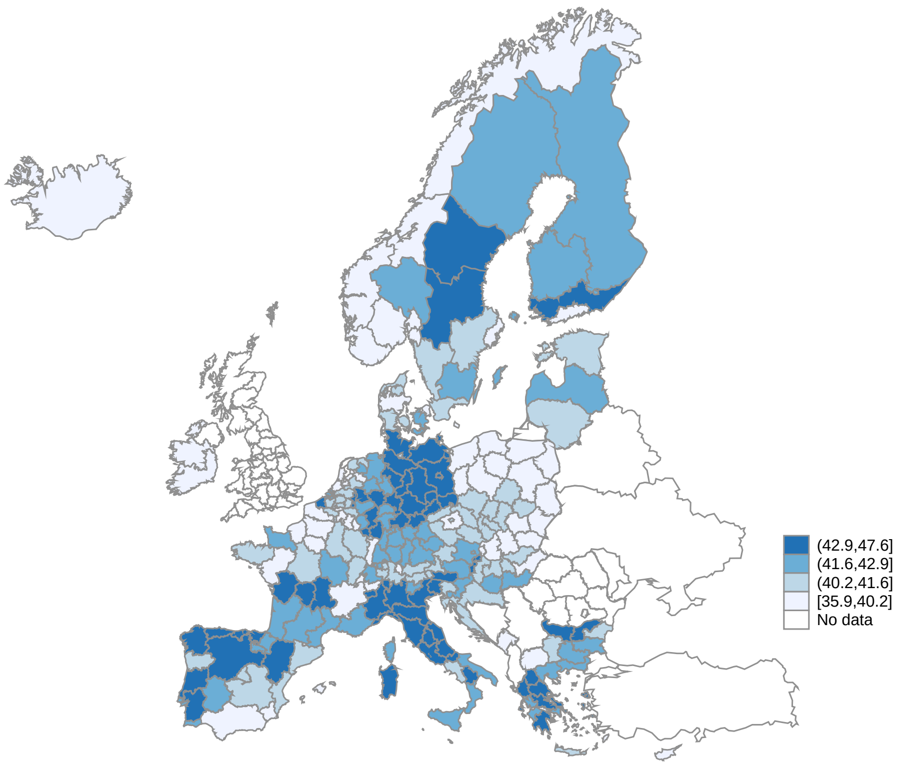
will become irrelevant in the globalized and highly connected world due to decreasing transport costs and advances in communication technologies. In their view, geography does not matter, which ultimately predicts a convergence across regions as a result of disappearing communication barriers. However, what we observe is exactly the opposite. Regional differences within countries are increasing, and location matters more than ever. In particular, forces are at play that are causing regions to be polarized. Therefore, understanding the determinants of regional disparities and potential reinforcement mechanisms is important for science, politics and society as a whole.

One aim of this dissertation is to consider the role of geographical migration for regional disparities. Migration is generally known to serve as an adjustment mechanism for territorial imbalances. Studies for the US show that wages in the US respond relatively flexible in response to adverse region-specific labour demand shocks (Blanchard and Katz, 1992). In particular, workers migrate from depressed regions to the better performing ones, thus equilibrating regional employment disparities. In contrast, interregional migration in Europe responds more slowly to a negative demand shock (Decressin and Fatás, 1995; Nahuis and Parikh, 2002), thus leading to strong and persistent regional employment disparities (Abraham, 1996; Mertens, 2002) as reflected in Figure 1. More importantly, such disparities may even be self-reinforcing if prosperous regions tend to attract predominantly young and high-skilled workers (Kanbur and Rapoport, 2005; Fratesi and Riggi, 2007). This may in turn trigger a cumulative process towards more polarized regions. This thesis therefore provides new insights into the determinants of migration selectivity, especially in a European context with strong and persistent employment disparities.

A further reason for a strengthening of regional inequalities, that has widely been neglected, are long-term demographic forces that are operating in the background of many advanced economies. In fact, the map of the most recent population demographics for European NUTS-2 regions shown in Figure 2 reveals a large geographical divide with respect to the population mean age. For instance, the first quintile (light blue) depicts the 20% regions with the lowest mean age in 2013, whose values range from 35.9 to 40.2. The fifth quintile (dark blue) contains the 20% regions with the lowest corresponding value. The maps show that whereas some European regions comprise a mean workforce age of up to 47.6 years, others show corresponding mean ages of only 35.9 years. Moreover, Eurostat projections over the next 50 years based on these figures predict that workforce ageing will continue in all European countries. This demographic trend has raised the concern that an ageing workforce may reduce productivity, innovative capability and thus, ultimately, competitiveness in the global, knowledge-based economy. More strikingly, workforce ageing is very likely to affect labour markets in very different ways on a regional scale. The reasons are closely linked to the fundamental forces of agglomeration and regional migration. In the case of Germany, urban areas are those that are very successful in attracting young and skilled workers due to their cultural amenities and career perspectives (Buch et al., 2014). On the other hand, rural areas that are suffering from depopulation and, in particular, from out-migration of their youngest and most educated workforce. Since the age of workers is known to be a key determinant of innovative behaviour, this demographic divide may likely turn into an innovation and, hence, economic divide. Although most present in the case of Germany, similar trends can be observed for other European countries, indicated by Figure 2.

A further aim of this thesis is therefore to document recent demographic trends for the case of Germany and to provide insights on the empirical relevance of agglomeration and regional polarization tendencies in this context. By focusing on Germany, we will be able to explore regional dynamics on a very local level for a fast ageing country that has the second

Figure 2: Population mean age for European NUTS-2 regions (2013)



Notes: Own illustration based on data from Eurostat.

highest median age worldwide behind Japan[1] and is characterised by a large demographic divide. Moreover, this thesis aims to contribute to the debate on the age-creativity link by assessing the causal impact of an ageing workforce on regional performance and by investigating potential complementarities and substitutabilities between different age groups. The thesis will show that the link of workforce ageing on regional performance must not necessarily be negative at the aggregate level due to externalities arising from knowledge interactions between different age groups. The idea is that fluid abilities (speed of problem-solving and abstract reasoning) are known to decrease at older ages, whereas crystallized abilities (ability to use skills, knowledge and experience) remain at high functional levels until late in life. It will be tested whether young and older workers are complementary in the production of knowledge and how this complementarities may enhance innovation through knowledge exchange at the level of local labour markets. To the authors' knowledge, such an investigation has not been conducted yet.

The first part of the thesis consists of 3 separate research articles (chapters) that were written

[1] See `http://esa.un.org/unpd/wpp/Documentation/pdf/WPP2012_HIGHLIGHTS.pdf`.

with co-authors. The papers are all based on employment register data (Beschäftigten-Historik – BeH) from the German Federal Employment Agency, an administrative data set that contains information on the population working in jobs that are subject to social insurance payments. For the different chapters, both the full BeH and 2% random samples are used. The BeH data is partly complemented with other regional data discussed in the data descriptions of the corresponding chapters. **Chapter (1)** lays the foundation for the subsequent chapters by exploring the spatial and temporal patterns of knowledge production and demographic measures within an Exploratory Space-Time Data Analysis (ESTDA). For the analysis, demographic measures including the average age and age dispersion as well as the share of creative professionals as one of the most important drivers of regional innovations are constructed. The workforce data are complemented by rich data from the European Patent Office (EPO) that include published patents in Germany. It is then explored to what extent innovations as well as creative and young workers tend to be geographically concentrated and how these concentrations have been evolving over time. Besides commonly used tools for cluster analyses, newer visualisation methods are applied that allow investigating the space-time dynamics of the spatial distributions and help to detect a potential reinforcement of clusters and spatial polarization tendencies. Moreover, the persistence of clusters as well as spatial contagion forces are investigated by means of transition probabilities. Overall, the results speak in favour of a demographic polarization trend across German regions. In particular, the detected post-reunification East-West divide is increasingly turning into a rural-urban divide. Whereas most urban regions are increasingly shaping a young and age-diverse workforce, almost all rural regions in both Eastern and Western Germany are affected by out-migration of their youngest cohorts. Since most East German regions constitute rural areas, they have been affected most by this trend, thus transitioning towards an age-homogenous economy with less mobile and older workers. However, the results also reveal a decent catching-up process of a few Eastern regions around the recently agglomerating capital city and a few other economic beacons that have increased their innovation output despite an ageing workforce. The findings indicate a first setting in of agglomeration forces after the transition from a communist to a market system. Due to the detected spatial contagion forces and cluster-wise path dependencies, regions are unlikely to reverse the trend though.

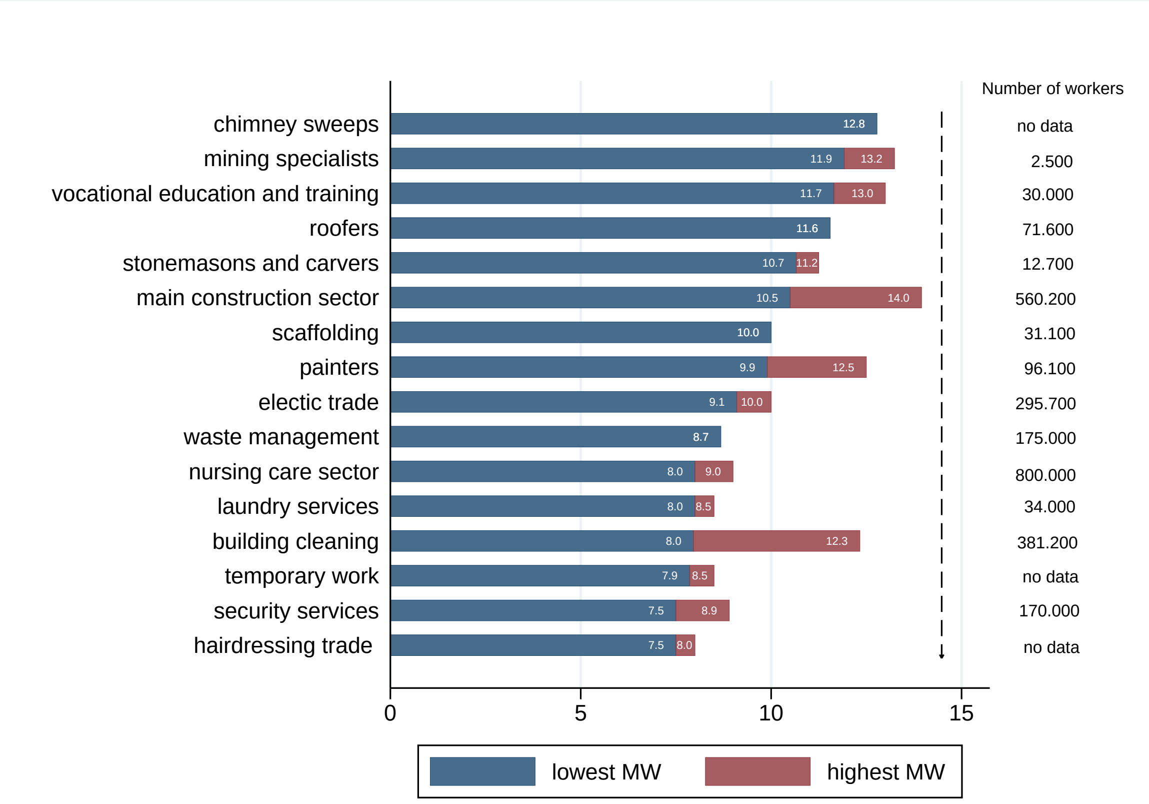
Motivated by the large demographic divide revealed in Chapter (1), **Chapter (2)** builds upon this descriptive preparatory work to examine the causal link between workforce age structure and patenting activity on the level of local labour markets. It also assesses potential complementarities and substitutabilities between different age groups within the local economy. The level of regional labour markets thereby constitutes the preferred unit of analysis for such an investigation, since the regional context appears to be most relevant for between-firm knowledge externalities and the generation of ideas (Peri, 2005). In order to address the potential endogeneity arising from e.g. selective migration, an Instrumental Variables (IV) approach is applied. In a first step, the age-creativity link is estimated in a quadratic specification, as is commonly done in the literature, using both cross-sectional and panel data. In a second step, a more flexible Translog production function is estimated using the number of young, middle-aged and older workers to gain insights into the complementarities and substitutabilities between these input factors. Overall, the results suggest a more complex pattern compared to the hump-shaped age-innovation profile typically found by existing studies. In particular, the findings indicate that younger workers boost regional innovations, but that this effect partly hinges on the presence of older workers in the same region. Moreover, cross-partial derivatives from Translog production functions suggest that abilities of younger workers and the experience of older workers are complementary in the production of knowledge. Despite this positive indirect effect of older workers on the production of knowledge, however, the findings point towards a reduced innovation level if demographic ageing shrinks the size of the younger workforce considerably in the future. Finally, Chapter (2) provides evidence that the difference in the age structure of the least and most innovative regions in Germany explains around a sixth of the gap in innovative performance, thus demonstrating the importance of demographic forces in shaping regional disparities.

One of the reinforcing mechanisms of regional disparities are skill-selective migration flows, as discussed above. Yet, the determinants of human-capital allocations across regions are still not fully understood. One of the few existing theoretical frameworks proposed by Borjas et al. (1992) links selective migration to wage disparities only. In particular, high-skilled workers ceteris paribus should be attracted to regions that best reward their abilities by paying high wage returns to their skills as reflected in a high wage inequality. However, modelling the migration decision as a wage-maximising process only may not be sufficient in a European context with less flexible wages and where regional differences are more driven by persistent employment disparities. For this reason, **Chapter (3)** extends the Borjas framework to allow for a selection mechanism based on both wage and employment differentials and models the average skill level of a migration flow as a positive function of both the wage and employment inequality in the destination as compared to the origin region. Unlike the Borjas framework, the model suggests that, besides mean wage, employment differentials also induce a positive skill sorting. The predictions for the average skill level of gross labour flows between German regions are then tested by regressing the average skill level of migration flows on the mean and dispersion of the regional wage and employment distribution. For the analysis, the panel dimension of the data is exploited in order to control for average time constant utility differentials between regions (e.g. amenity differentials). Moreover, to account for the endogeneity, a Difference Generalized Method of Moments (GMM) estimator proposed by Arellano and Bond (1991) is applied. The findings confirm the relevance of regional employment disparities for skill-selective migration, while regional wage differentials have no robust and significant impact. This chapter thus fills an important gap in understanding the self-reinforcing nature of interregional employment disparities. Although focusing on interregional migration, the main findings should apply to cross-country migration as well. Still, the results suggest that the recent emergence of intra-European migration flows from Southern Europe towards high employment countries such as Germany that has been found by Bertoli et al. (2013) is likely skill-biased, thus potentially aggravating the current North-South divide depicted in Figure 1.

Part II: Minimum Wage Effects Along the Wage Distribution

Whereas Part I deals with the influence of geography and demographic forces, Part II looks at the role of wage-setting institutions for labour market inequality. Investigating the influence of institutional factors has become increasingly relevant due to rising wage inequality (as measured by the 90-10 log wage differential) in many industrialized countries during the last four decades (Katz and David, 1999; Machin and Van Reenen, 2008). Although the wage structure in a few countries such as Germany stayed remarkably stable until the beginning of the 1990s it then also started to increase in these countries too. For instance, several studies show that wage dispersion increased both at the top and at the bottom of the wage distribution in Germany (for an overview see Fitzenberger, 2012). Compared to other international findings, the increase in German wage inequality is economically relevant and has led to decreasing real wages (Antonczyk et al., 2010). Particular attention has been given to the increase in lower tail wage inequality and the rise of the low-wage sector.

Figure 3: Minimum wage regulations in Germany as of August 2014 (in Euros)



Notes: Own illustration based on data from Destatis. The numbers of workers refer to all workers subject to social security contributions and are taken from the Confederation of German Trade Unions (DGB).

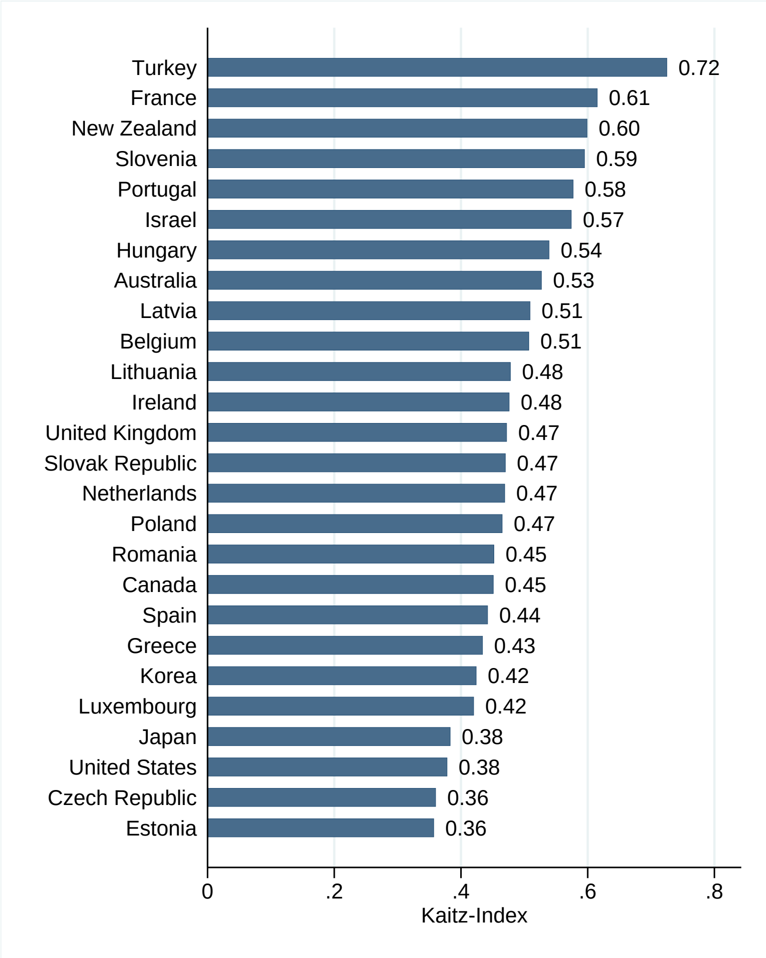
Rising wage inequality is one reason why most industrialized countries have implemented MWs targeted at increasing the wages of the working poor. The level and design of MW regulations, however, varies considerably between countries. In Germany, the first MW was introduced at a sectoral level for the construction sector in January 1997. Companies faced increasing competition from Eastern European firms who offered their services in Germany very cheaply due to lower wages that even undercut the collective wage agreements in Germany. In order to protect these traditional crafts, trade unions and employers associations agreed as part of a general collective bargaining agreement on the implementation of a MW. However, since not all subsectors agreed on these regulations, only the main construction sector and a few months later other subsectors including the roofing, painting, varnishing and electric trade industry ended up implementing the first sector-specific MWs in Germany. For many years these industries were the only ones with a MW. Meanwhile, 16 industries have implemented a MW (see Figure 3 for an overview of existing MW regulations in Germany). The levels vary between Eastern and Western Germany as well as between worker groups (e.g. skilled vs. unskilled), captured by the 'highest' and 'lowest' MW in Figure 3. Most recently, the newly elected grand coalition decided to implement a statutory MW of 8.50 Euros, beginning in 2015. An extensive discussion on the pros and cons of MWs predated this development in the public media.

There are only a few studies on MW effects in Germany. Most of these studies have recently come out of a comprehensive policy evaluation project in 2011 on behalf of the Federal Ministry of Labour and Social Affairs (BMAS) where six German research institutions evaluated eight out of twelfth sector-specific MWs in Germany. The aim of the evaluation was to assess the impact of the MW on employment, worker protection and competition in the sectors under study. Among others, the Centre for European Economic Research (ZEW) analysed the MW effects in the German roofing sector, which this part of the thesis builds upon (for a detailed report see Aretz, 2011). A summary of the results is published in Aretz et al. (2012b).

Two research articles in this part of the thesis originate from this policy evaluation project, but extend this work to highlight some new stylized facts on MW spillovers. Most MW research focuses on the average employment and average wage effects on workers with a binding MW for whom the wage falls below the MW level. However, depending on the production technology, the MW may also affect workers whose wage lies above the MW level. Such MW spillovers are mostly neglected in the empirical literature and many studies even use non-binding workers as a counterfactual for Difference-in-Differences analyses. The aim of the two research articles in Part II is to demonstrate how a MW may affect the employment probabilities and earnings of those workers located in the upper part of the wage distribution that are typically assumed to be unaffected by such policy interventions.

The German roofing sector turned out to be an interesting sector of study in this respect. First, the MW level in the roofing sector is one of the highest for unskilled workers in Germany (compare Figure 3). In fact, only chimney sweeps, mining specialists and workers conducting vocational education and training services are subject to higher MWs (introduced only recently). Moreover, since average wages are particularly low in the roofing sector, the MW bites particularly hard in this sector as reflected by a Kaitz-Index (the ratio of the MW level and the median wage) of around 1 in Eastern Germany. Compared to other OECD countries, the bite has to be considered exceptionally hard also by international standards. In fact, according to Figure 4, the Kaitz-Index varies between 0.4-0.7 for OECD countries. The German roofing sector thus represents an ideal setting to study MW spillovers since its bite is likely to aggravate wage and

Figure 4: Minimum wage relative to median wages for OECD countries in 2012



Notes: Data taken from OECD data base. Median wages are calculated for full-time workers.

employment effects along the entire wage distribution.

One research article presented in this part of the thesis is written with co-authors and another is single-authored. Both are based on two administrative linked employer-employee panels, one of which contains the full sample of workers in the roofing sector over the observation period of interest. Both investigations exploit a quasi-experiment since, for institutional reasons, the MW was introduced only in parts of the construction sector, including the roofing sector. Uncovered, yet comparable, sub-sectors may thus serve as a benchmark for the counterfactual development in the roofing sector in order to derive the treatment effects of the MW on subsequent outcomes. In particular, **Chapter (4)** investigates the causal impact of the MW on the probability of remaining employed in the roofing sector. Since the entire construction sector experienced a dramatic decline in demand after the end of the unification boom in the mid 1990s that almost halved the workforce in Eastern Germany, this is a highly relevant employment outcome. Results from an inter-sectoral and intra-sectoral comparison are contradicted, providing first insights into potential employment spillovers in the sector. The latter two identification strategies are then applied to a comparison of workers with and without a binding MW across sectors, which makes it possible to directly estimate the employment effects along the entire wage distribution. Unobserved heterogeneity at the individual level is thereby taken into account, which may be relevant if employers mainly substitute workers along unobservable skills. The findings indicate that the average probability of remaining employed in the roofing sector has deteriorated due to the MW introduction, especially in Eastern Germany where the bite of the MW was particularly hard. Moreover, the results from the comparison of workers with and without a binding MW suggest negative employment outcomes for East German workers located higher up in the wage distribution. According to interviews with sector insiders, capital-labour substitution seems to be important in driving this finding. Overall, this chapter highlights the need for a broader perspective on the employment impact of MWs and also put doubts on any attempt to identify employment effects of MWs by comparing workers with and without a binding MW within a covered sector.

Chapter (5) complements Chapter (4) by focusing on the MW effects on the earnings (rather than on employment) along the distribution. The descriptive assessment of the hourly earnings distribution throughout the project already hinted at a strong wage compression not only at the lower but also at the upper part of the distribution, thus suggesting negative wage spillovers on high-wage earners. Motivated by this interesting observation, a similar quasi-experiment as in Chapter (4) is exploited using the wage distribution of uncovered sub-sectors as a counterfactual for the earnings of roofers in the absence of the policy reform. For the analysis, a recently developed unconditional quantile regression method suggested by Firpo et al. (2009) is applied that allows investigating the MW effects at each quantile of the distribution while keeping other factors constant. To yield further insights into a potential within-group (workers with similar characteristics) rather than overall wage compression effect, the results are contrasted to findings from conditional quantile regressions. Overall, the mean impact of the MW seems to miss a lot. In particular, the results suggest significant real wage increases of lower-decile workers that ripple up to the 0.6th quantile in Eastern Germany, whereas the weaker wage effects in Western Germany (5% at the lower tail) pillar up to about the median worker. However, the estimates also reveal some unexpected side effects of the reform. According to the estimates, the MW caused a reduction in real wages by up to 5% in Eastern Germany for the highest quantiles that mostly comprise skilled and experienced workers. The wage compression effect thereby not only reflects lower entry wages, but rather indicates wage restraints among upper quantile workers. This finding for Eastern Germany is consistent with evidence in favour of a rising cost burden. Particularly smaller firms with limited influence on market prices (price-takes) and fewer possibilities for substituting labour by capital have limited the scope for wage increases among their skilled employees. However, these wage restraints among highly paid workers only became possible due to increasing numbers of high-skilled workers queuing for jobs, which has strengthened the bargaining power of firms vis-a-vis skilled workers they still employ, a finding that indirectly results from the work presented in Chapter (4). As a consequence, wage differentiation and thus incentives for human capital investments have been shrinking in the sector and might also explain labour shortages that firms have been facing recently as reported by sector insiders.

Part I

Demographic Change and Regional Labour Market Disparities

1 Demographic Ageing and the Polarization of Regions - An Exploratory Space-Time Analysis[1]

Joint with Roberto Patuelli[2]

Abstract: Demographic ageing is expected to affect labour markets in very different ways on a regional scale. Contributing to this debate, we explore the spatio-temporal patterns of recent distributional changes in the worker age structure and innovation output for German regions by conducting an Exploratory Space-Time Data Analysis (ESTDA). Besides commonly used tools, we apply newly developed approaches which allow investigating joint dynamics of the spatial distributions. Overall, we find that innovation hubs tend to be located in areas with high skill concentrations, but also seem to coincide with favourable demographic age structures. We show that these concentrations are persistent over time due to clusterwise path dependence and spatial contagion forces. The spatio-temporal patterns speak in favour of a demographic polarization process of German regions where the post-reunification East-West divide is increasingly turning into a rural-urban divide.

Keywords: innovation, demographic ageing, exploratory space-time analysis, regional disparities

JEL-Classification: J11, R12, R23

[1]Forthcoming in *Environment and Planning A*.
[2]University of Bologna and Rimini Centre for Economic Analysis (RCEA).

1.1 Introduction

Demographic ageing has increasingly become one of the most pressing challenges that industrialized economies are facing in the 21st century. According to the latest Eurostat projections for the next 50 years, workforce ageing will continue in all European countries, though its magnitude, speed and timing are likely to vary. Demographic trends have raised the concern that an ageing workforce may increase existing regional disparities in a global, knowledge-based economy, as innovation potential is likely to depend on the age structure of local workers. One reason is that innovation activity is strongly concentrated in agglomerated areas due to advantages derived from externalities due to the colocation of specific industries (localisation economies) and the accessibility of firms to a variety of skilled workers (urbanisation economies). Particularly young and educated workers are attracted to urban areas and thick labour markets due to cultural amenities and better career opportunities (Moretti, 2011; Buch et al., 2014). This may further increase existing disparities, as urban areas that are already good at attracting human capital and good jobs tend to attract even more. In contrast, rural areas suffering from depopulation further diminish their human capital base (brain drain).

Whereas such divergence process has been demonstrated for US labour markets (Moretti, 2012), such phenomena are less clear for the German and, more generally, European markets with limited worker and firm mobility relative to the US.[3] Moreover, Germany is an interesting case of study due to its strongly ageing workforce and a large demographic divide. Spatial imbalances may have been reinforced by increasing labour-force participation of women who, seeking job opportunities, have increasingly been moving to prosperous urban areas. This trend has affected fertility patterns across regions and may further aggravate the rural-urban divide. Eastern regions have been suffering strongly from an ageing workforce due to age- and gender-selective out-migration. However, understanding the role of agglomeration forces in triggering a self-reinforcing process towards polarization might be interesting beyond the German case. For instance, Puga (2002) provides a discussion, based on location theories, of the possible (negative) causes of polarization within European countries, highlighting, for example, the role

[3] Südekum (2008) investigates the spatial variation of human capital across West German regions for a historical time period showing that concentration forces are much lower compared to the US. However, he focuses only on qualification degrees and it is unclear how these forces are developing lately, especially as Germany is moving towards an aging knowledge-based economy.

of transport infrastructure.

The objective of this paper is to contribute to this debate and explore the spatial and temporal patterns of knowledge production and demographic measures by means of an Exploratory Space-Time Data Analysis (ESTDA) to yield insights into recent demographic trends and how they may change the regional innovation landscape. There are several studies that have already explored the spatial distributions of economic performance or income across European regions using local and global measures of spatial association (Le Gallo and Ertur, 2003; Ertur and Koch, 2006; Dall'erba, 2005; Patacchini and Rice, 2007). However, these studies use more general indicators of economic performance and consider space-time dynamics only partially. Exceptions are a study by Le Gallo (2004) and more recent studies by Hierro et al. (2013) and Fazio and Lavecchia (2013), which deal with the persistence of regional disparities by exploiting spatial transition probabilities. We build on this literature and extend these approaches by newer visualisation methods for a comprehensive ESTDA of our innovation and demographic measures.

In particular, our contribution is fourfold. Firstly, we describe the spatial distribution of regional age structure, human capital and innovation performance in the interesting case of a fastly ageing Germany and discuss the possible theoretical links between these variables. We thereby do not focus only on the average age of workers, but also consider age diversity in order to capture a wider picture of the workforce age distribution. Secondly, we use a rich data set from the European Patent Office (EPO) that includes all published patents in Germany. By focusing on patents as one direct measure of the innovation process at the regional level, we are better able to capture innovativeness than more general indicators of economic performance such as productivity and economic growth. Furthermore, by including the share of creative professionals in our analyses we additionally explore one of the most important drivers of regional innovation (Florida, 2002; Florida and Gates, 2003). Thirdly, instead of only using static (spatial) methods such as Local Indicators of Spatial Association (LISA), we apply new visualization tools such as directional Moran scatterplots, developed by Rey et al. (2011), which allow investigating the space-time dynamics of spatial distributions, and help to detect a potential reinforcement of clustering and polarization. In addition, we calculate LISA transition probabilities as suggested by Rey (2001) to study the persistence of regional disparities. To our knowledge, this paper is the first to combine all the above methods and provide a comprehensive ESTDA on the themes of labour force ageing and innovation output. Finally, the paper contributes to the discussion on

East-West convergence (divergence) after reunification which, after almost 25 years, still has important consequences for the theory and the design of policies in the demographic context.

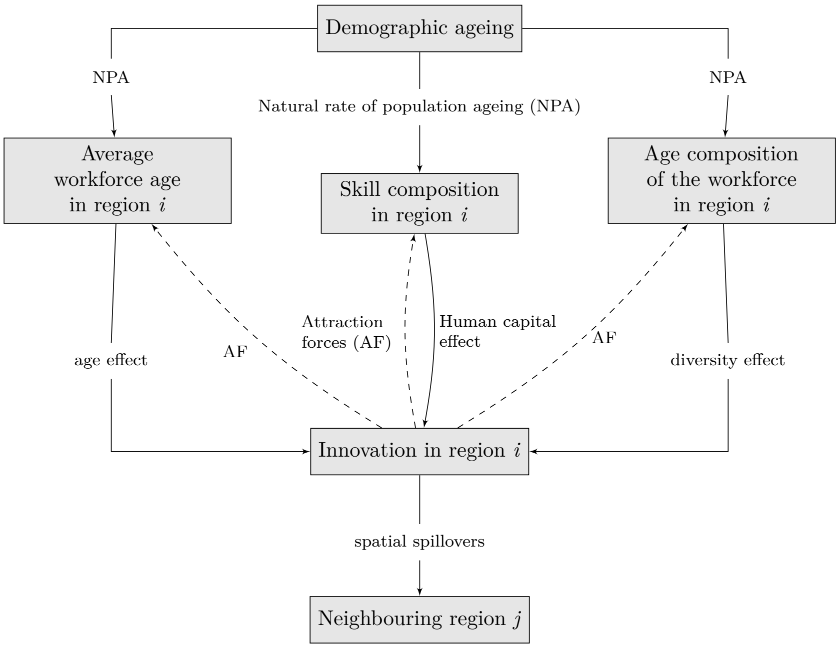
Overall, our study shows that location matters in an aging knowledge-based economy as suggested by the detected spatial concentrations. In particular, we find that innovation hubs tend to be located in areas with high skill concentrations, but also seem to coincide with favourable demographic age structures. The study further demonstrates the persistence of these concentrations over time as indicated by clusterwise path dependence and spatial contagion forces in shaping the distributions. Moreover, the spatio-temporal patterns speak in favour of a demographic polarization process of German regions. According to our results, the post-reunification East-West divide is increasingly turning into a rural-urban divide. Whereas most urban regions are increasingly shaping a young and age-diverse workforce, most rural regions in both East and West Germany are affected by out-migration of their youngest cohorts.

The paper is structured as follows. First, we discuss the potential theoretical mechanisms of demographic polarization trends before introducing the database in Section 1.3. In Section 1.4 we then conduct an ESDA by first testing for spatial randomness (*global* Moran' I) and describing patterns of spatial clusters and outliers (*local* Moran' I). In Section 1.5 we then move to the space-time dynamics by investigating changes in the spatial clusters over time (directional Moran Scatterplots) and path-dependencies (LISA Transition matrices). Finally, Section 1.6 concludes.

1.2 The Polarization of Regions

There are several channels through which demographic ageing may affect the competitiveness of regions and thus, ultimately, increase regional disparities. As Figure 1.1 illustrates, the natural rate of population ageing does not only change the average age, but also shapes the age composition of the regional workforce through regionally differentiated fertility rates, changes in average life expectancy as well as differences in labour force participation rates. This may have several consequences for the production of knowledge in a region. On the one hand, an increasing average age may have a diminishing effect on the creativity of workers (age effect), which is known to decline with age (Simonton, 1988; Bratsberg et al., 2003; Jones, 2010). On the other hand, regions with increasing cohorts of older workers may even profit from an ageing workforce

Figure 1.1: Channels through which demographic ageing may trigger a trend towards more polarized regions



(diversity effect) if they develop a favourable age composition. The reason is that older workers are endowed with specific experience and (tacit) knowledge that may be complementary to the one of younger workers, as shown by a recent study at the regional level by Arntz and Gregory (2014). Moreover, there might be an indirect effect of demographic ageing on regional innovation through changes in the human capital base (human capital effect) arising, for instance, from older and formally more skilled worker cohorts (especially in the East) retiring and younger workers entering the labour market. The fact that the human capital base is an important driver of knowledge production and regional growth has been stated in various research (Lucas, 1988; Florida, 2002; Florida and Gates, 2003). The overall impact of these channels on regional knowledge production is, however, far from clear. The few existing studies at the country or regional level suggest a negative impact on GDP growth (Lindh and Malmberg, 1999; Brunow and Hirte, 2006) and total factor productivity (Feyrer, 2008). In contrast, Arntz and Gregory (2014), who use patent counts and citations, show that the overall impact of workforce ageing must not necessarily be negative, once the endogeneity of regional workforces is controlled for.

Given the overall effect, agglomeration forces then set in and reinforce existing disparities by pushing firms and skilled labour towards the innovation hubs, while leading to depopulation and brain drain in rural areas. The reasons for such spatial agglomerations are higher productivity and wages (Glaeser et al., 1992; Rauch, 1993; Ciccone and Hall, 1996) arising from local externalities (localization and urbanization economies). Skilled, young and more mobile workers in particular are attracted to urban areas, which offer several advantages such as cultural amenities and better career perspectives (Moretti, 2011; Buch et al., 2014). Note that such in-migrants may originate from both other regions within the country and from abroad.[4] Spatial clustering in knowledge production and human capital then occurs through localized knowledge spillovers (Jaffe, 1989; Audretsch and Feldman, 2004) and population relocation to surrounding urban counties (suburbanization). Overall, the above mechanisms may trigger a cumulative process towards increased polarization. This hypothesis is supported by several studies in the migration literature that show how selective migration, induced by interregional differences in wage and employment opportunities (Arntz et al., 2014) may lead to increasing spatial inequalities (Kanbur and Rapoport, 2005; Fratesi and Riggi, 2007) rather than serving as a re-equilibrating mechanism. In the next sections we explore the spatial distributions of innovation and demographic measures in the case of Germany in order to provide insights on the empirical relevance of such agglomeration and regional polarization tendencies.

1.3 Innovation and Demographic Measures

The present study focuses on workforce rather than population data, since we assume the regional age structure to affect regional innovation mainly through the working rather than overall population. For the calculation of the workforce age structure, we exploit the regional file of the Sample of Integrated Labour Market Biographies (SIAB) from the Institute of Employment Research (IAB) for the years 1995-2008. The data set is an employment subsample provided by the German Federal Employment Agency and contains information on workers that are subject to social insurance contributions by their employers, thus excluding civil servants and self-employed individuals. The data includes individual employment histories on a daily basis and contains, among others, information on the age and occupations of workers. We use annual cross sections at the cut-off date of 30 June and calculate regional indicators of demographic

[4] A study by Poot (2008) discusses several reasons why immigration may affect regional competitiveness in the context of demographic ageing.

composition including average age, age dispersion (standard deviation) and the share of creative professionals[5] (which we will refer to as our human capital base or skills). We restrict the analysis to full-time employed individuals subject to the social insurance contribution, that is, excluding minors and unemployed workers.[6] Furthermore, we restrict our data set to working individuals between 18 and 65 years of age to avoid selection problems that would be associated, for instance, to the fact that those few underage workers are undergoing a vocational training. As a regional definition, we use the 332 labour market regions defined in the regional file of SIAB data. These regions reflect aggregated counties to the extent that they comprise at least 100,000 inhabitants. We focus on this detailed regional level instead of using more aggregated labour market regions in order to distinguish between different degrees of urbanization.

As a measure of regional innovativeness, we use patent data which are provided by the European Patent Office (EPO). The use of such direct outcome measures is still rare in the literature dealing with the effects of ageing workers on competitiveness, especially in regional level studies[7], but should be able to better capture innovativeness than more general indicators of economic performance. Our data set contains patent data both at the applicant and inventor level. Whereas the applicant (quite often, the firm) is the holder of the patent right, the inventors are the actual inventors cited in the document. We focus on patent inventors, since we are interested in their spatial distribution rather than in the location of the formal holder, which is often one of the firm's headquarters. Since patents may have been developed by serval inventors located in different regions, we apply a fractional counting approach to assign to every region the respective share of the patent. For instance, an inventor who developed a patent in Mannheim with one further individual working abroad would generate 0.5 patents for this region. Following this procedure for each of the 332 regions, we calculate the number of patent applications for the years 1995-2008. Since the number of inventions of a region may simply reflect its size rather than the knowledge production efficiency, we furthermore condition the number of patents (multiplied by 100) by the number of employed workers of the region, obtaining a measure of patent production per 100 workers.

There are several advantages and disadvantages of using patenting data at the regional level (Giese and von Reinhard Stoutz, 1998; Giese, 2002). On the one hand, patent applications

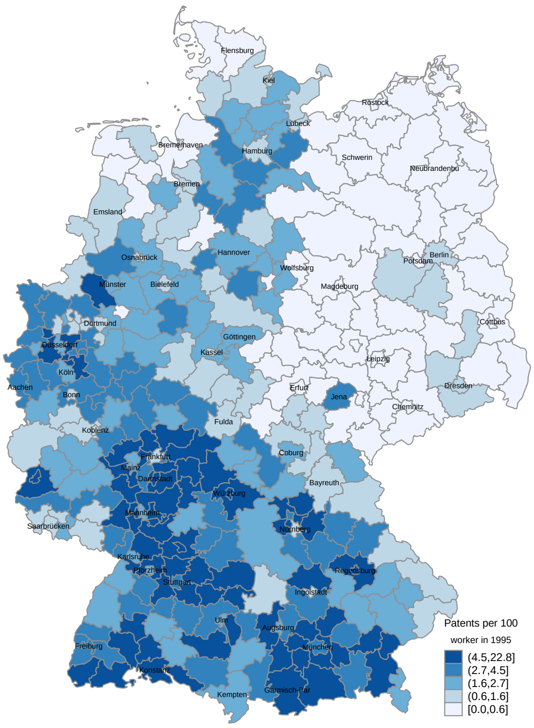
[5]For the classification of creative professionals, we follow Möller and Tubadji (2009).
[6]We hypothesize here that part-time workers are equally employed across regions.
[7]See, for instance, Brunow and Hirte (2006), Feyrer (2008) and Lindh and Malmberg (1999).

are a useful indicator of research and invention activities at the local level, as they include information on the origin of inventor activities, that is, the place of residence and therefore, indirectly, the approximate location of the research institute. On the other hand, not every invention becomes the subject of a patent application, nor does a patent necessarily become a marketable product or process. Moreover, the reasons for a patent application may not only rest on protecting an invention against unjustified use, but may reflect strategic concerns such as securing and extending regional markets, prestige advertisement and the demonstration of innovative capacity to economic counterparts (e.g., shareholders or funding partners). Despite these disadvantages, empirical evidence by Acs et al. (2002), who provide an exploratory and a regression-based comparison of the innovation count data and data on patent counts at the lowest possible levels of geographical aggregation, suggests that patents provide a fairly reliable measure of innovative activity. Similarly, a survey study by Griliches (1998) concludes that patents are a good indicator of differences in inventive activity.

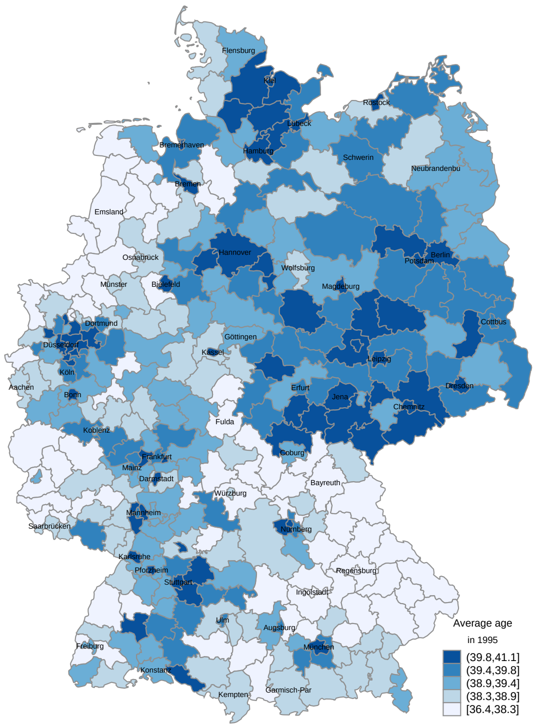
Figure 1.2 maps patents per 100 worker and average workforce age for our 332 regions for the initial year as well as the absolute change between 1995 and 2008. Similar maps for workforce age dispersion and the share of professionals are shown in Appendix 1.A.1. For instance, the first quintile in Figure 1.2a (light blue) depicts the first 20% of least innovative regions in 1995, whose values range from 0.2 to 1.4 patents per 100 workers. The fifth quintile (dark blue) contains the 20% most innovative regions. The maps show that innovations are mostly generated in urban counties around West German cities such as Duesseldorf, Frankfurt, Stuttgart, Freiburg, Nueremberg and Munich. These regions also employ most creative professionals. In contrast, only a few East German major cities such as Jena were halfway competitive in the production of knowledge a few years after reunification. The spatial distribution of average age further reveals that only a few West German regions exhibit relatively old workforce, including major cities and urban counties around Kiel, Hamburg, Bremen, Hannover, Duesseldorf, Frankfurt, Stuttgart, Nueremberg and Munich. The fact that most rural regions in West Germany comprise relatively young and age-diverse workforces reflects historically large shares of conservative farming families with traditional role models that led to relatively high fertility rates, particularly in Bavarian and North-Western counties (around Emsland). In contrast, East German regions depict relatively old and homogenous workforces indicating that plant closures and out-migration of young workers after reunification strongly affected their age structure. The latter has already

Figure 1.2: Regional quantile maps for innovation and average workforce age for the initial year 1995 and absolute changes between 1995-2008

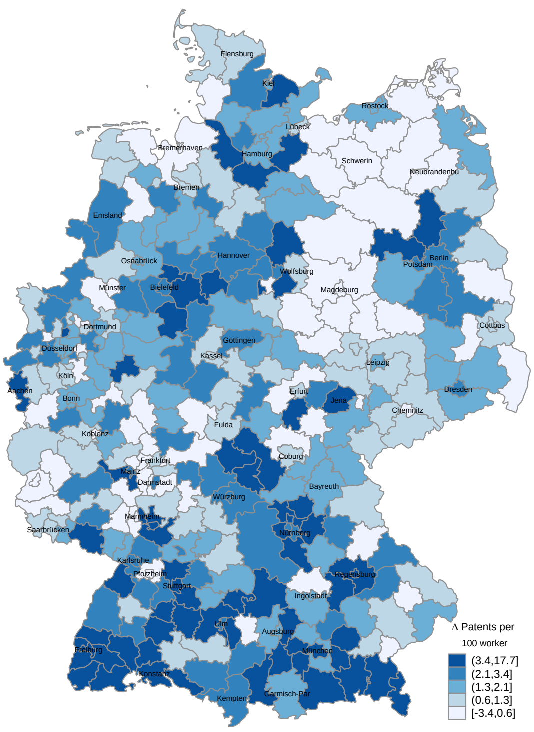
(a) Patents per 100 workers in 1995



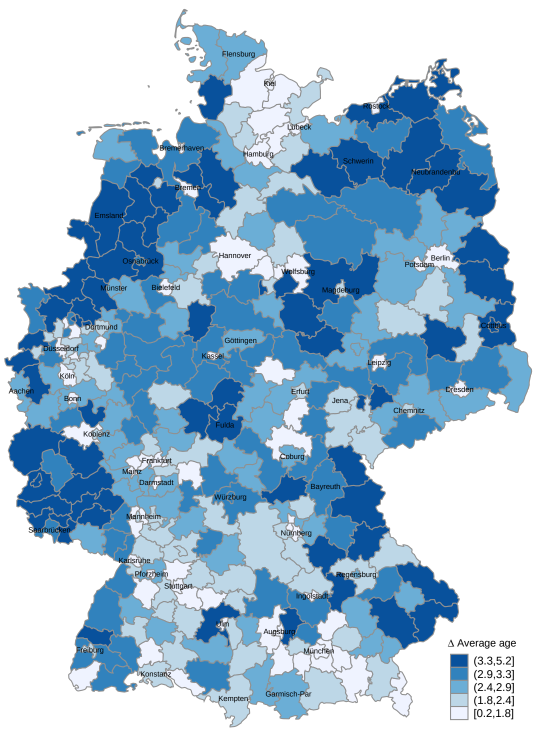
(b) Average workforce age in 1995



(c) Δ Patents per 100 workers



(d) Δ Average workforce age



been confirmed by Burda and Hunt (2001) and Hunt (2004), who study the years between reunification and the millennium and find that East-West migrants tend to be young and better educated compared with stayers.

However, looking at changes over time (during the 14-year period considered here) suggests that the East-West divide is turning into a rural-urban divide (see Appendix 1.A.2 for a map of German regions by agglomeration status). Whereas urban areas were able to hold their average age constant during the last 14 years, rural regions in both parts of the country experienced a strong demographic ageing process, with increases in their average age up to 5.2 years. These developments also reflect rising labour force participation of women during the last 30 to 40 years. In particular, young and qualified women increasingly find better career perspectives in urban areas, thus depressing the fertility rates of rural regions (in addition to depopulation). This is particularly true for East Germany that comprises many rural counties. Overall, these findings already indicate spatial dependencies in the ageing and innovation processes, as many regions seem to have experienced a similar trend to their surrounding regions. As discussed in Section 1.2, this might reflect agglomeration forces that are reenforcing existing spatial inequalities and which may lead to a polarization of regions. In order to shed light on these deep economic forces, we thus explore spatial regimes and investigate the space-time dynamics of the spatial distributions in the next section. Such an analysis enables the detection a potential reinforcement of clusters and spatial polarization tendencies.

1.4 Global and Local Spatial Autocorrelation

In the present section, we test the hypothesis of spatial randomness using the *global* Moran's I (MI) statistic and use Local Indicators of Spatial Association (LISA) to visualize *local* patterns of spatial associations (clusters). We conduct the static analysis for the initial year 1995, a few years after reunification, as a benchmark for the following analysis. In Section 1.5, we then analyse the space-time dynamics of the observed spatial clusters (or outliers) across the period 1995-2008. The latter will also allow to reveal potential distributional shifts in Eastern Germany due to the transition from a communist system to a market economy and where agglomeration forces might have started to set in and shape the spatial patterns of our innovation and demographic measures.

1.4.1 Global spatial autocorrleation

Since the distribution of workers cannot be expected to be random in space, we test for global spatial autocorrelation using the MI indicator, which provides a single summary statistic describing the degree of clustering present in spatial data. In particular, it allows implications on whether, for instance, highly (lowly) innovative regions are often surrounded by regions that are also highly (lowly) innovative. This is interesting, since it reflects agglomeration forces and spatial spillovers. Moreover, it allows to classify whether a region is part of a relevant cluster, such as a hot (cold) spot, or rather an outlier. Note that this information can be used in any regression analysis as a proxy for e.g. knowledge spillovers between regions.

We first define the structure of the spatial relationship by considering a spatial weights matrix based on rook contiguity that assumes neighbouring relationships between regions by shared borders.[8] The spatial weights matrix provides information on the spatial proximity between each pair of locations i and j, while the diagonal values of the weights matrix are set to zero. We standardize the matrix so that the elements of each row sum to one (row-standardization).[9] We define the spatial lag of a variable y_i in region i as the average value of a variable evaluated at its neighbouring units. We then construct a bivariate scatterplot with standardized values y_i on the horizontal axis and their spatial lags $\sum_{j=1}^{N} \widetilde{W}_{ij} y_i$ on the vertical axis (Moran Scatterplot, see Figure 1.3). As a covariance and correlation measure we consider the Moran's I statistic, which constitutes a measure of the overall spatial dependence.[10] The MI can be interpreted as a regression coefficient resulting from the regression of the spatial lag $\widetilde{W}_{ij} y_i$ on y_i (Anselin, 1996). Values of I greater (smaller) than $E(I)$ indicate positive (negative) spatial autocorrelation.

Figure 1.3 shows the Moran Scatterplots for the demographic and innovation measures. Each of the points in Figure 1.3 represents a combination of a regions' value in 1995 and its corresponding spatial lag. The values on the x- and y-axes are standardized so that the vertical and horizontal lines represent the national values and divide the scatterplot into 4 quadrants

[8] As recently shown in the literature (e.g. see Patuelli et al. 2012), the choice of the spatial weights matrix is often of little importance, since different geography-based matrices tend to have strongly correlated weights. In a regression framework, multiple matrices may be tested ex post, for example by means of Bayesian model comparison (LeSage and Pace, 2009).

[9] The elements of the standardized weights matrix are defined as follows: $\widetilde{W}_{ij} = W_{ij} / \sum_{j=1}^{N} W_{ij}$, where $W_{ij} = 1$ if i and j are defined as neighbours and $W_{ij} = 0$ if otherwise.

[10] We define the MI as follows: $I = N/S_o(\sum_{i=1}^{n} \sum_{j=1}^{n} \widetilde{W}_{ij}(y_i - \bar{y})(y_j - \bar{y}))(\sum_{i=1}^{n}(y_i - \bar{y})^2)$ for $i \neq j$, where $S_o = \sum_{i=1}^{n} \sum_{j=1}^{n} \widetilde{W}_{ij}$.

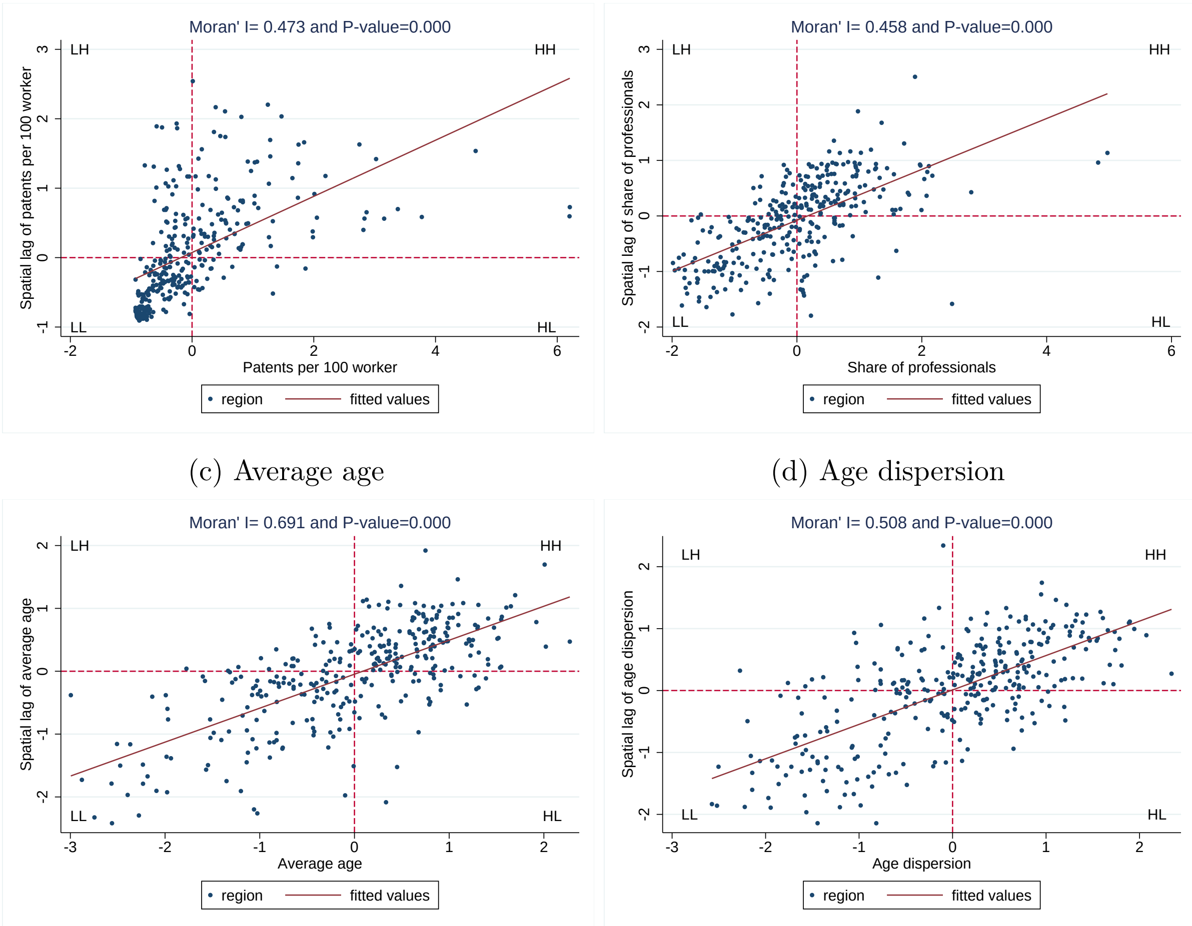
Figure 1.3: Moran's I scatterplot for patents per 100 workers, average age, age dispersion and share of professionals in 1995

(a) Patents per 100 workers

(b) Share of professionals

(c) Average age

(d) Age dispersion



that correspond to the following four different types of spatial association (anticlockwise from top right): high-high (HH), low-high (LH), low-low (LL) and high-low (HL). For instance, a HH region exhibits a high number of patents per worker and is surrounded by regions that exhibit a high number of patents as well. Both HH (hot spots) and LL (cold spots) represent regimes of positive spatial association, whereas LH and HL indicate negative association. The calculated MI for global autocorrelation is represented by the slope of the line interpolating all points in the scatterplot since it is based on standardized values.

Figure 1.3 shows for all variables significant degrees of spatial autocorrelation. Most regions are either in the first (top right) or third quadrant (bottom left). Note that the last row in the Appendix 1.A.3 summarizes the total amount of regions in each quadrant. For instance, in the case of patents per worker, almost 30% of all regions (98 out of 332) fall into the first quadrant and 50% in the third. Interestingly, the points agglomerate dominantly in the third quadrant and become more dispersed with increasing values. This result indicates large clusters of scarcely productive regions, whereas clusters of highly productive regions seem rare. Compared to the US, for instance, the concentration of high-tech industries thus seems less. A clearer indication of clustering is found for average age, for which positive spatial association appears to be wide, in terms of both higher and lower values. According to Column 5 in Table 1.A.3, 40% of regions fall into the first quadrant and a similar fraction into the third. The pattern is similar and stronger for age dispersion and the share of creative professionals, thus indicating spatial concentrations of young and diverse workers in creative occupations. These observed patterns are statistically significant according to the MI coefficients, which are all above zero.

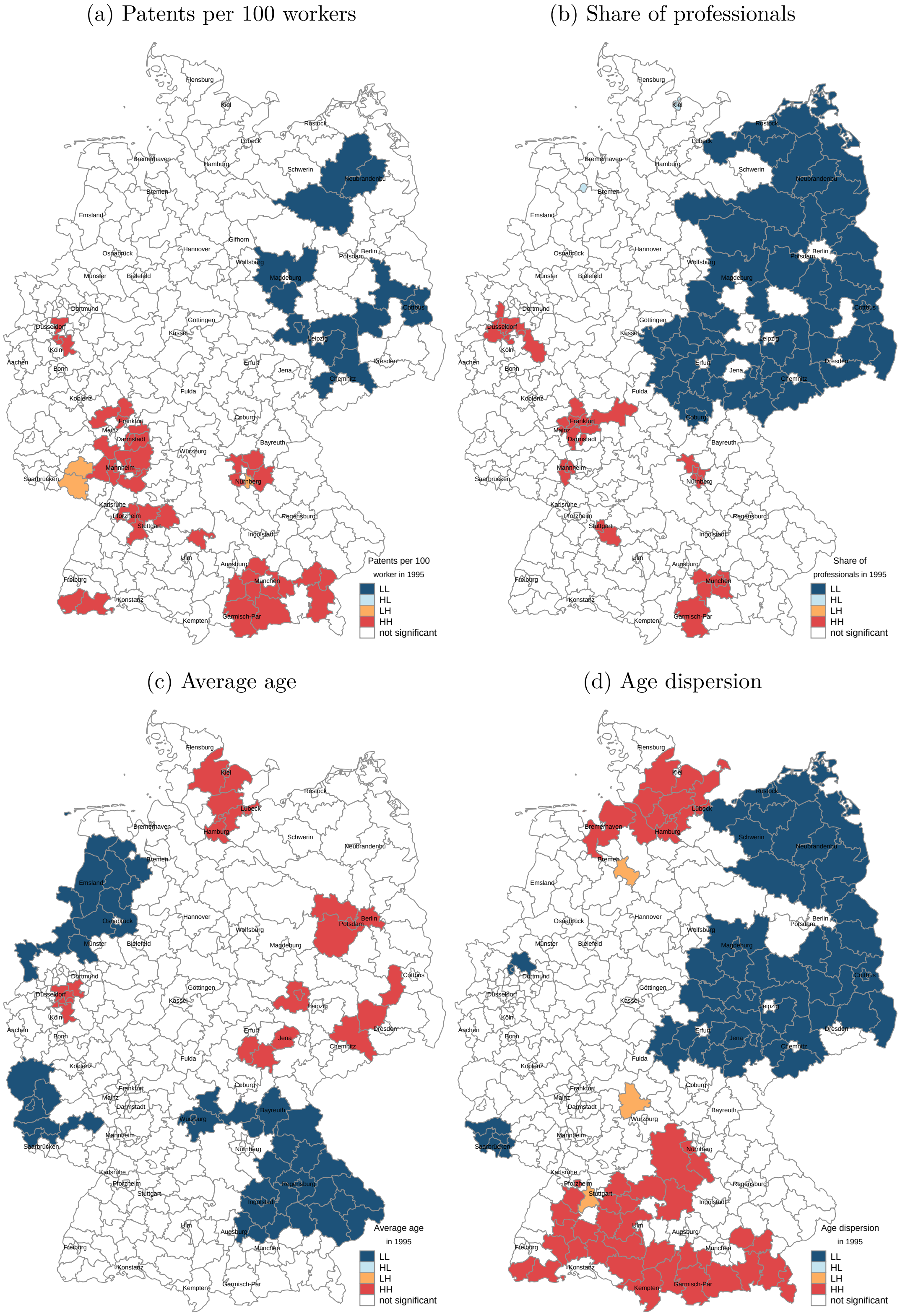
1.4.2 Local Indicators of Spatial Association

In the present section, we aim to locate the observed clusters and assess their spatial extent. Since these questions cannot be answered by means of global measures of spatial autocorrelation, we use Local Indicators of Spatial Association (LISA) as proposed by Anselin (1995). The local version of MI gives an indication on the significance of local spatial clustering for each region.[11] Similarly to the global MI statistic, significance can be determined through the expected value and variance. The interpretation is similar. A positive LISA indicates clustering of HH or LL values in and around i, whereas a negative LISA indicates a spatial outlier, that is either HL or LH.

Figure 1.4 shows the LISA cluster maps (again for the initial year 1995) for our four variables and where only values that are significant at the 5% level are presented. The maps reveal a large East-West divide. In particular, they show large clusters of lowly innovative regions in rural and sparsely populated counties in East Germany around Neubrandenburg, Magdeburg, Leipzip, Chemnitz and Cottbus. In contrast, the innovation hubs are located in mostly urban counties in Western and Southern Germany around Duesseldorf, Frankfurt, Mannheim, Stuttgart, Freiburg, Nueremberg and Munich. There is almost no significant outlier, indicating that regions are unlikely to be a high (low) innovative region in a low (high) innovative cluster. Moreover, these high-tech clusters coincide with spatial concentrations of creative professionals. According to

[11] We define the local MI as follows: $I_i = (\sum_{j=1}^{N} \widetilde{W}_{ij}(y_i - \bar{y})(y_j - \bar{y}))(\frac{1}{N}\sum_{j=1}^{n}(y_i - \bar{y}))$.

Figure 1.4: LISA cluster maps for patents per 100 workers, average age, age dispersion and share of professionals in 1995



our contingency tables displayed in Appendix 1.A.3, 80% of all (significant and insignificant) regions in an innovation hub coincide with professional worker hotspots. In contrast, about 65% of all low innovation clusters coincide with low skill concentrations.

Looking at the LISA cluster maps for average age shows surprisingly few significant clusters of old regions in Eastern Germany. Particulary rural regions still seem to profit from historically high fertility rates. After reunification, the almost entire Eastern workforce constitutes one large cluster of age-homogenous regions. Considering West Germany, there are only two old-age clusters in the Ruhr district, which has been struggling with its structural change and around Kiel in Northern Germany, whereas almost all Bavarian regions in Southern Germany and regions around Emsland in North-Western Germany show young worker concentrations. The latter reflect areas of prosperous growth with a specialisation in the agricultural sector. Clusters of age-diverse regions are mostly located in Southern and Northern Germany and reflect regions with a relatively balanced mix between young and older cohorts. Many of these regions are also part of an innovation hub as indicated by the contingency tables.

Overall, the investigation of local and global spatial autocorrelation underlines the importance of spatial dependencies. In particular, LISA analyses indicate that innovation hubs tend to be located in areas with high skill concentrations, as one would expect, but also seem to coincide with a favourable demographic age structure. One explanation may be that high-tech clusters tend to be very successful in attracting young workers and shaping an age-diverse workforce by keeping the older and more experienced ones in the labour force. As discussed in Section 1.2, this might in turn increase the innovation potential of these regions even further. In fact, agglomeration externalities seem to have not reached their upper bound yet, as indicated by limited concentrations of innovators. In contrast, we observe strong negative clustering of lowly inventive and unattractive areas in Eastern Germany that are facing considerable difficulties in coping with the transition to the innovation sector and that are lacking a (creative) human capital base.[12] These regions are particularly old and homogenous age-wise. However, since the patterns describe the initial situation after reunification, it remains to be shown how these clusters (and the large East-West divide) developed during the observed 14-year period and whether the existing disparities have lead to a spatial polarization trend as agglomeration theories

[12]This problem is exacerbated by the devaluation of the preexisting workers qualifications, especially for white-collar ones (e.g., in management, administration), which makes evaluating human capital in the former East Germany highly challenging.

would suggest.

1.5 Space-Time Dynamics

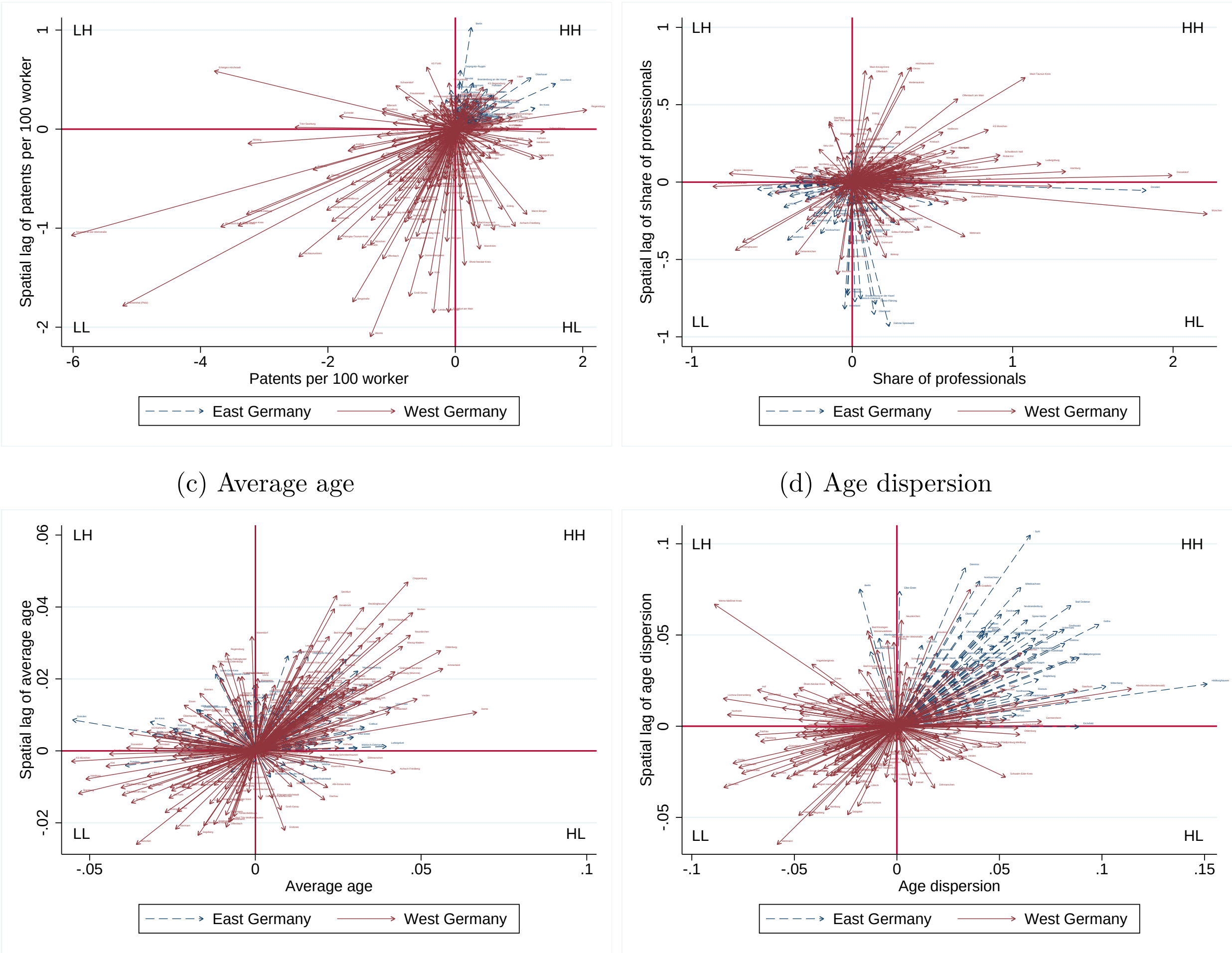
So far, we have gained insights into the spatial dimension of the data distributions, measured by the values of the initial year 1995. We are now interested in how such distributions evolved over time and whether there are any observable trends. Furthermore, we investigate the stability of the observed spatial patterns over time to reveal potential path dependencies. Most studies analysing the evolution of a variable's spatial distribution visually compare different geographical maps for separate points in time. Such approaches make it very difficult to analyse relative movements across time and space. For this reason, we apply new methods that are designed to address this limitation.

1.5.1 Standardised Directional Moran Scatterplots

In this section, we investigate regional dynamics using Standardized Directional Moran Scatterplots (SDMS, Rey et al. 2011). For each variable, we calculate Moran Scatterplots for the years 1995 and 2008 separately (as described in Section 1.4) using relative values (to the national value). Note, that this time period is particularly interesting, due to the large second wave of selective migrants that moved from East to West Germany during the end of the 1990s, as discussed, for instance, by Arntz et al. (2014). This wave of migrants had its peak in 2001 and is expected to have changed the regional distribution of the workforce age structure. We plot each region's value in 1995 and 2008 into the same Scatterplot and connect both points to receive directional movement vectors.[13] All vectors are normalized by the 1995 national value to produce the SDMS shown in Figure 1.5. Whereas the arrowheads point to the regions' 2008 relative value, the vectors' starting point (at the origin) represents its corresponding value in 1995. The SDMS thus captures how a regions' position and spatial association developed between 1995 and 2008 relative to the national trend. For instance, a move of a region towards the first (HH) or third quadrant (LL) reflects the strengthening/emergence of positive spatial clustering (in a meliorative and worsening perspective, respectively) or the inversion of a previous

[13] We also checked the plots using the 3-year averages for the periods 1995-1997 and 2006-2008 as connecting points. Since the picture did not change much, we are confident that our results are not driven by exceptional occurrences in the years 1995 and 2008.

Figure 1.5: Standardized Directional Moran Scatterplots for patents per 100 workers, average age, age dispersion and share of professionals (1995 to 2008)



(a) Patents per 100 workers

(b) Share of professionals

(c) Average age

(d) Age dispersion

opposite position (e.g., an LL region moving towards the origin, i.e., improving, will have a HH movement). On the other hand, movements towards the second (LH) or fourth (HL) quadrant reflect negative clustering tendencies (i.e., a local divergence process). The longer the movement vector, the larger the relative movement compared to the mean. As a robustness check we also calculated the average values for the periods 1995-1997 and 2006-2008 as alternative connecting points. Since the results did not change much, we stick to the former.

Figure 1.5 shows the SDMS for our four variables.[14] Movements of East German regions are shown in blue (dashed line), and West German regions are represented by in red (solid line). For patent production, the figures show large movements towards cold spots (movements

[14]For improved readability, we dropped Neustadt an der Weinstrasse and Frankenthal in graph (a) and Frankfurt am Main and Berlin in graph (d), since there target points lie far right and far left of all other regions.

towards the third quadrant). Among these are mostly urban counties around the West German metropolitan cities Frankfurt, Darmstadt and Ludwigshafen (e.g. Neustadt an der Weinstrasse, Darmstadt-Dieburg, Main-Taunus-Kreis), although some of these regions are moving from high initial values. In contrast, we observe only small relative movements towards innovation hot spots (movements towards the first quadrant) that are dominated by Berlin's peripheral regions Overhavel, Havelland, Barnim and Potsdam city. Also, Jena and its surrounding urban counties Weimar and Ilm improved in terms of creating high-tech clusters. These regions are increasingly successful in creating knowledge-based industrial districts. Their favourable developments are also reflected by their increases in the share of professional workers, despite the low average values of their surrounding neighbours (reflected by movements towards the fourth quadrant). Apart from these exceptional developments, East Germany as a whole is experiencing negative clustering tendencies with respect to creative professional shares, whereas almost all West German regions show higher skill concentrations. These findings suggest that, despite a small catching-up process for a few agglomerated areas, the overall degree of patenting and its geographical concentration remains low in Eastern Germany. In particular, Eastern regions are increasingly facing difficulties in speeding up the accumulation of professional skills, despite their improvements in innovation. This may also help explaining why we find surprisingly limited evidence for positive co-movements of innovation and human capital concentration (see Appendix 1.A.4).

Compared to our innovation measures, the SDMSs for the demographic variables reveal much stronger polarization tendencies. The dominant movements towards the first quadrant for average age are driven mostly by rural areas both in East and West Germany because of rural-urban migration of particularly young workers. In contrast, regions moving towards the third quadrant include major West German cities such as Munich, Stuttgart, Frankfurt, Hamburg. The latter developments may reflect suburbanization processes for which the existing cluster slowly spread beyond administrative borders to new regions. The results indicate an increasing concentration of young workers in cities, whereas rural regions are suffering from depopulation and ageing workers. Of course, since most peripheral regions are located in Eastern Germany (see Appendix 1.A.2), this development mirrors the former divide between the two parts of the country. However, the concentration of younger workers have not coincided much with high-tech clustering, as suggested by our contingency tables. This is somehow surprising, and may reflect still low innovative regions (such as the Berlin suburbs) that have been relatively improving their innovation output, despite an ageing workforce. On the other hand, the figures also indicate that 44 (mostly West German) regions have been moving both towards lowly innovative and high-age clusters. The association between innovation and age dispersion is much clearer. In particular, an increasing clustering of age-homogenous regions (mostly rural regions in Eastern Germany) coincides with lower clustering of high-tech industries, whereas geographical environments with an increasingly age-diverse workers base (mostly young cities in Western Germany) seem to be quite successful in the generation of knowledge (spillovers).

Compared to the initial state (after reunification), where the almost entire Eastern economy reflected low innovation performance and ageing workers (relative to the West), the recent trends speak more in favour of a rural-urban divide. Whereas most urban regions are increasingly shaping a young and age-diverse workforce, almost all rural regions in both East and West Germany are affected by out-migration of their youngest cohorts. Since most Eastern regions constitute rural regions, they have been affected most by this trend, thus transitioning towards a age-homogenous economy with less mobile older workers. However, the results also reveal a decent catching up process of few Eastern regions around the recently agglomerating capital city and few other economic beacons in the East that have increased their innovation output, despite an ageing workforce. These findings may indicate preliminary evidence of agglomeration forces (suburbanization processes) setting in after the economic transition and changing the spatial distribution slightly. Whether regions are able to turn the trend around or will rather remain in their state due to geographical contagion is something we will investigate by means of transition probabilities in the following section.

1.5.2 Space-Time Transitions

In this section we calculate LISA transition matrices in order to track the evolution of the investigated variables from a spatial clustering perspective. The method is based on the classical Markov chain approach, which allows to study time dynamics between different groups (e.g., quantiles). From a methodological viewpoint, the proposed LISA transition matrices are obtained similarly to the standard probability transition matrices. We follow Rey (2001) and investigate the transitions of regions between the four different types of spatial association outlined above

Table 1.1: LISA transition probabilities for innovation and demographic measures (1995-2008)

Variable	Cluster/ Outlier in period t	Cluster/Outlier in period $t+1$ $(HH)_{t+1}$ (1)	$(LH)_{t+1}$ (2)	$(LL)_{t+1}$ (3)	$(HL)_{t+1}$ (4)	Initial shares in 1995 (5)	Steady state (6)
Patents per 100 worker	$(HH)_t$	90.3	6.3	0.5	2.9	29.5	28.3
	$(LH)_t$	12.1	79.6	8.1	0.2	15.4	15.6
	$(LL)_t$	0.4	2.6	93.6	3.4	49.4	48.1
	$(HL)_t$	8.4	1.9	20.7	68.9	5.7	8
Share of professionals	$(HH)_t$	94.9	4.2	0.0	1.0	14.5	16.9
	$(LH)_t$	2.8	0.1	95.0	2.1	17.8	24.5
	$(LL)_t$	0.0	1.1	97.2	1.6	55.4	46.9
	$(HL)_t$	1.4	0.0	6.8	91.8	12.4	11.7
Average age	$(HH)_t$	91.4	4.6	0.6	3.5	44.6	38.3
	$(LH)_t$	16.5	68.1	14.2	1.3	9.0	11.8
	$(LL)_t$	0.3	4.9	88.6	6.3	35.5	36.7
	$(HL)_t$	9.6	1.8	17.2	71.4	10.8	13.3
Age dispersion	$(HH)_t$	82.7	7.6	8.5	1.2	45.2	29.5
	$(LH)_t$	17.6	63.6	15.0	3.8	13.3	13.5
	$(LL)_t$	1.3	5.0	85.6	8.1	29.8	39.2
	$(HL)_t$	12.4	2.4	18.3	66.9	11.8	17.8

(HH, LH, LL, HL) to allow a quantitative assessment of contagion effects. For a detailed technical explanation see Appendix 1.A.5. The calculated transitions are shown in columns (1) to (4) in Table 1.1. Column (5) includes the share of regions in the different states at the beginning of the period, whereas Column (6) corresponds to the computed steady state shares (expected long-run equilibrium shares). For instance, the probability of a highly innovative region surrounded by highly innovative regions (HH) to remain in its current state over each time period (a year) is 90.3% (see row 1 and column 1), whereas the probability of remaining a LL region accounts to 93.6% on average (row 3 and column 3). All variables show fairly high off-diagonal probabilities. In particular, age dispersion shows relatively high transition probabilities reflecting high dynamics over time. The two transitions with the highest off-diagonal probabilities are generally those for regions moving from LH to HH and from HL to LL (negative contagion), that is, transitions where a region is 'infected' by the state of its neighbours (Hierro et al., 2013). This result thus indicates that it is highly likely, for an outlier, to become part of its surrounding cluster speaking in favour of strong contagion forces at place.

The probability of negative contagion is thereby higher compared to positive contagion, that is, it is more likely for an HL region to become LL than for an LH region to become HH. The only exception is the share of professional workers. This result stands in contrast to the one of

Hierro et al. (2013) who stress that positive spatial contagion (transitioning from LH to HH) is more likely to be expected than negative contagion (from HL to LL). Finally, movements from LH to LL and from HL to HH seem fairly high as well in the case of average age and age dispersion. Obviously, single regions may also end up pulling their neighbours up/down to their status, but the probability of this occurrence is much lower. Thus, the probability to reverse the trend, for underperforming regions with an old and homogenous age structure, for instance in East Germany, is low. There appears then to exist a clusterwise 'path dependence', where it is not only a region's own history that influences its chances of modifying its status quo in the future. In fact, its surrounding environment plays a role in limiting the range of possible future outcomes, or in favouring different outcomes, when there is a mismatch between a region's state and the one of its surrounding areas. In particular, geographical areas facing a downward trend in terms of innovation performance and an ageing population are likely to pull other regions down as well. One reason for such geographical contagion forces might be age-selective outmigration, which has an impact on the demographic structure of neighbouring regions, as well as social networks and interregional ties. Clearly, such hypothesis may be further tested by means of regression modelling.

Furthermore, columns (5) and (6) of Table 1.1 show the initial shares of clusters/outliers in 1995 and the long-run ones suggested by our data, by means of the ergodic steady-state distributions. We find that most regions were either LL or HH in the initial period and this appears to be true in the long run as well. For instance, 29.5 (49.4)% of regions were in a highly (lowly) innovative cluster in 1995 and the expected share in the long run is 28.3 (48.1)%. The probability of being part of a cluster - both highly and lowly innovative - is therefore expected to decrease. The latter finding may reflect the slow recent catching-up process of Eastern regions. Our results further suggest an increasing likelihood of staying or becoming a region with high professional worker shares or being located near such a cluster (enlargement of both HH and LH) whereas the opposite is true for low skill concentrations (shrinking of both HL and LL). Moreover, our demographic measures indicate that LL clusters will become wider, with a decrease in the size/number of the HH clusters. Put differently, an increasing share of regions will comprise a homogenous workforce which reflects the rural to urban migration of young workers and the depopulation of peripheral regions, particularly in Eastern Germany. This might also explain the increasing concentrations of young workers (in urban areas) as suggested by the steady state values.

Overall, the transition matrices presented above suggest that location matters for the evolution of regional innovation and of the workforce characteristics, in the sense that the evolution of a region depends strongly on its neighbouring regions. In particular, it is unlikely for a region to reverse its trend in a highly interdependent geographical environment, indicating clusterwise path dependence. Moreover, the evidence suggests that outlier regions face a high probability to become part of the surrounding cluster due to strong contagion forces.

1.6 Conclusion

This paper contributes to the debate on demographic change in Europe and the potential role of spatial dependencies and agglomeration forces in triggering a cumulative process towards more polarized regions. In particular, we explore the spatio-temporal dynamics of regional innovation output, workers demographics and the creative human capital base for Germany. We apply newly developed approaches in order to detect spatial regimes or other forms of spatial heterogeneity for the investigated variables as well as its spatio-temporal dynamics.

The detected spatial concentrations suggest that location matters strongly in the German context. In particular, we find that innovation hubs tend to be located in areas with high skill concentrations in Western and Southern Germany, but also seem to coincide with favourable demographic age structures. In contrast, we observe strong negative clustering of lowly inventive regions with ageing working populations in Eastern Germany. These regions are still transitioning to a knowledge-based economy and are attempting to build up a human capital base. Transition probabilities indicate that these concentrations are likely to remain relatively stable due to a strong clusterwise path dependence as well as contagion forces in shaping the spatial distributions. Hence, it is not only a region's own history that influences its chances of modifying its status quo in the future, but also the surrounding environment that plays a role in limiting (or favouring) the range of possible future outcomes.

Temporal changes in the spatial concentrations further suggest that the former East-West divide is increasingly turning into a rural-urban divide. Whereas most urban regions are increasingly shaping a young and age-diverse workforce, almost all rural regions in both East and West Germany are affected by out-migration of their younger cohorts. Since most Eastern regions are rural, they have been affected most by this trend, thus transitioning towards age-homogenous economies with less mobile older workers. However, the results also reveal a catching up process of few Eastern regions around the recently agglomerating Berlin and a few other economic beacons in the East that have increased their innovation output, despite an ageing workforce. The findings might indicate first attempts of agglomeration forces setting in after moving from a communist system to a market economy. Our results are somewhat in contrast to other studies such as Südekum (2008) who explores historical data and finds that spatial concentrations in West Germany are not big enough to trigger a self-reinforcing spatial concentration. Looking at a broader set of variables for a more recent time period, our findings seem to suggest quite the opposite for the innovation sector in an aging economy where agglomeration and urbanisation seem to matter stronger. Overall, we can not confirm Friedman (2005)'s hypothesis of a "flat world", according to which location will become irrelevant in the globalized and highly connected world due to decreasing transport costs and advances in communication technologies.

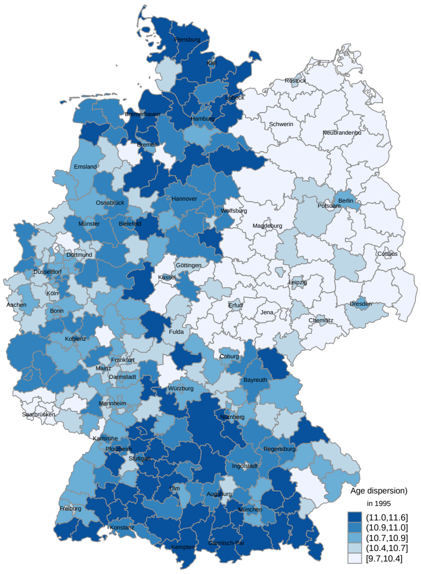
Our results have several policy implications. First, local policymakers aiming at reducing spatial inequalities should take into account the role of agglomeration and contagion forces in the innovation process, as well as (sub-)urbanisation trends in affecting workforce dynamics of spatially contiguous areas. In particular, major cities are gaining importance for young and skilled workers because of thick labour markets and rich amenities (Moretti, 2011; Buch et al., 2014). Due to spatial contagion, regions are unlikely to reverse this trend. Rather, urban regions that are already successful in attracting a young and diverse human-capital base appear to further attract such workers and aggravate a positive feedback loop process. From a national perspective, promoting innovation activities in beacon regions (for instance in Eastern German) with the aim of exploiting knowledge spillovers and (positive) agglomeration externalities might thus be more promising for the economy than turning around the trend in depopulating and less attractive rural areas. Widespread promotion of Eastern regions as done by the joint Federal Government/Länder scheme for "Improving regional economic structures (GRW)" could be revisited in this regard. At the same time, regional policy strategies might consider "big push" type of policies to move the region to a good equilibrium (Moretti, 2011; Kline, 2010) and trigger a self-reinforcing process in the positive direction. Furthermore, regions should cooperate more with other neighbouring regions in shaping an attractive metropolitan area for young workers, rather than competing against them. This might include alliances in the education system such as different universities with different specialisations, but which are complementary.

Our findings also have implications for future work in this field. In particular, our analysis may serve as a departure point for any analysis trying to measure the impact of demographic ageing on firm or regional performance. In particular, the causal relation between workforce ageing and a regions' innovation potential is still far from understood. Recent attempts by Arntz and Gregory (2014) move forward in this direction and explore the benefits of a more age-diverse workforce. The presence of strong clustering in the demographic variables, and of very specific outliers with regard to innovation, further suggests that spatial econometric techniques may be exploited when investigating such research question.

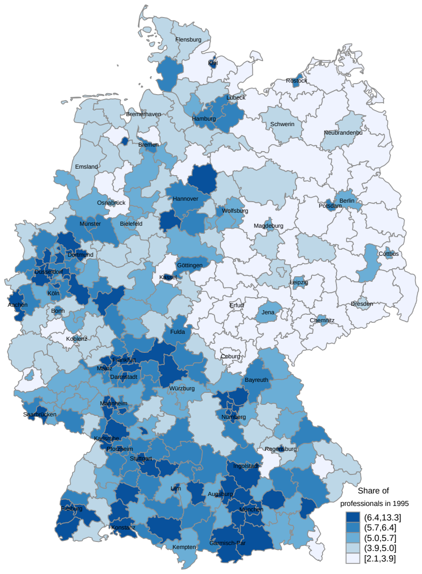
1.A Appendix

1.A.1 Regional quantile maps for workforce age dispersion and professional shares for the initial year 1995 and absolute changes between 1995-2008

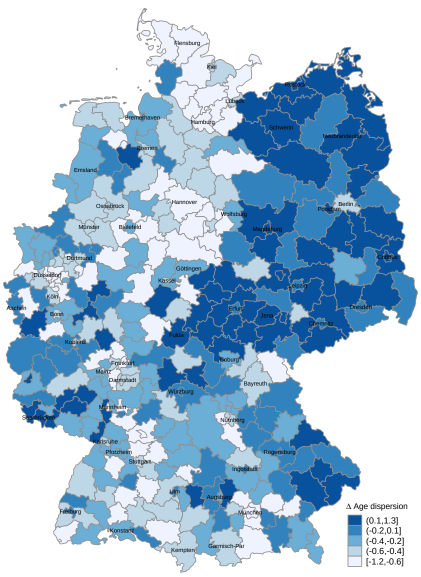
(a) Workforce age dispersion in 1995



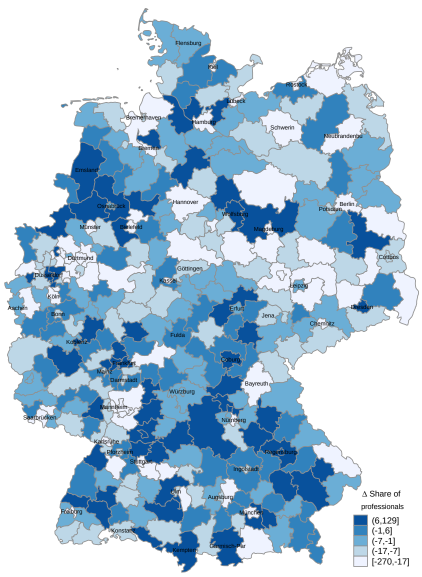
(b) Share of creative professionals in 1995



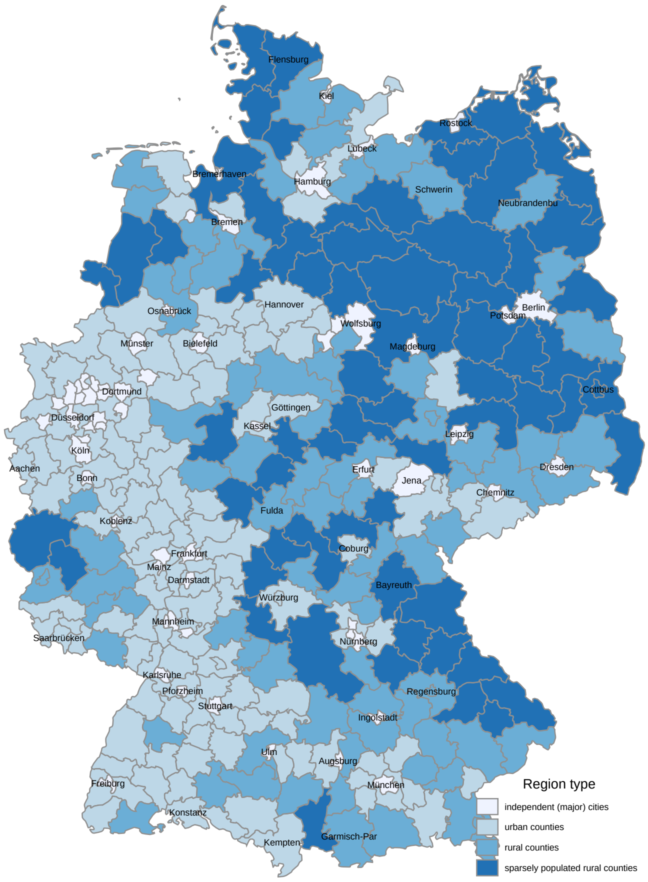
(c) Δ Workforce age dispersion



(d) Δ Share of creative professionals



1.A.2 Classification of German regions according to their agglomeration status



1.A.3 Contingency tables for LISA cluster maps in Figure 1.4

Variable	Cluster/Outlier type	Patents per 100 workers				
		High-High (1)	Low-High (2)	Low-Low (3)	High-Low (4)	Total obs. (5)
Average age	High-High	37	13	94	4	148
	Low-High	10	3	16	1	30
	Low-Low	39	25	42	12	118
	High-Low	12	10	12	2	36
		Pearson χ^2 = 76.4687 Pr = 0.000				
Age dispersion	High-High	59	27	51	13	150
	Low-High	18	10	15	1	44
	Low-Low	9	6	80	4	99
	High-Low	12	8	18	1	39
		Pearson χ^2 = 54.3235 Pr = 0.000				
Share of professionals	High-High	77	29	14	8	128
	Low-High	15	8	10	1	34
	Low-Low	5	7	108	7	127
	High-Low	1	7	32	3	43
		Pearson χ^2 = 10.3431 Pr = 0.323				
	Total obs.	98	51	160	23	332

Notes: The table reads as follows. The value in row (1) and column (1) indicates that 37 regions are both part of an innovation hot spot and old-age cluster, whereas the value in row (4) and column (1) means that 39 regions are part of a high-tech cluster and at the same time exhibit high concentrations of young workers.

1.A.4 Contingency tables between types of movements in the Standardized Directional Moran Scatterplots shown in Figure 1.5.1

Variable	Movements towards a Cluster/Outlier type	Patents per 100 workers				
		High-High (1)	Low-High (2)	Low-Low (3)	High-Low (4)	Total obs. (5)
Average age	High-High	64	30	44	24	162
	Low-High	27	14	17	15	73
	High-Low	5	5	10	7	27
	Low-Low	15	9	35	11	70
		Pearson χ^2 = 20.51 Pr = 0.015				
Age dispersion	High-High	71	20	33	15	139
	Low-High	9	8	18	18	53
	High-Low	5	9	18	6	38
	Low-Low	26	21	37	18	102
		Pearson χ^2 =44.65 Pr = 0.000				
Share of professionals	High-High	27	20	43	19	109
	Low-High	18	4	12	9	43
	High-Low	32	18	36	19	105
	Low-Low	34	16	15	10	75
		Pearson χ^2 = 15.87 Pr = 0.070				
	Total obs.	111	58	106	57	332

Notes: The table reads as follows. The value in row (1) and column (1) indicates 64 positive co-movements toward old-worker concentrations and high-innovation clusters, whereas row (4) and column (1) tell us that we observe 15 regions that move both in the direction of higher innovation clustering and young worker concentrations.

1.A.5 Calculation of LISA Transition Matrices

In order to compute LISA transition probabilities, we follow a markov chain approach. First, we specify a state probability vector $P_t = [p_{1t}, p_{2t}, p_{3t}, p_{4t}]$ that represents the probability of a region to be in one of the four states (in our case, the four quadrants of the Moran scatterplot) in period t, where $t = 1, 2, ..., 14$ in our case. We then define a 4×4 transition probability matrix, $\mathbf{M} = [m_{ij}]$, showing the likelihood of a region to remain in initial state i or to move from state i in period t to state j in period $t + 1$ during the 14-year period. Transition probabilities are assumed to be time-invariant, that is we assume a homogenous markov chain. Given these assumptions, the state probability vector in period t can be written as $P_t = P_0\mathbf{M}^t$, where P_0 is the initial state vector. In the long-run, the markov chain converges to the steady state vector d.

2 What Old Stagers Could Teach Us - Examining Age Complementarities in Regional Innovation Systems

Joint with Melanie Arntz[1]

Abstract: Concerns have been raised that demographic ageing may weaken the competitiveness of knowledge-based economies and increase regional disparities. The age-creativity link is however far from clear at the aggregate level. Contributing to this debate, we estimate the causal effect of the workforce age structure on patenting activities for local labour markets in Germany using a flexible knowledge production function and accounting for potential endogeneity of the regional workforce structure. Overall, our results suggest that younger workers boost regional innovations, but this effect partly hinges on the presence of older workers as younger and older workers turn out to be complements in the production of knowledge. With demographic ageing mainly increasing the older workforce and shrinking the younger one, our results imply that innovation levels in ageing societies may drop in the future. Moreover, differences in the regional age structure currently explain around a sixth of the innovation gap across German regions.

Keywords: regional innovation system, demographic ageing, knowledge production function, regional disparities, age complementarities

JEL-Classification: R12, R23, J11

[1]Universtity of Heidelberg and Centre for European Economic Research (ZEW), Mannheim.

2.1 Introduction

With accelerating demographic ageing in most industrialized economies, concerns have been raised that an ageing workforce may reduce creative performance and thus, ultimately, the competitiveness of the affected countries in the global, knowledge-based economy. These concerns are fueled by numerous studies on the creative performance of scientists and artists that suggest a peak productivity in middle-ages and a declining performance thereafter (Lehman, 1953; Simonton, 1988; Oster and Hamermesh, 1998; Bratsberg et al., 2003; Jones, 2010).[2] However, age-related declines in mental abilities must not necessarily translate into a diminishing innovative performance at the aggregate level if there are knowledge externalities between age-heterogenous individuals with complementary skills. The reason is that whereas younger workers are endowed with higher abilities in generating and recombining new knowledge, older workers tend to be have accumulated more experience and knowledge in how to use and apply existing skills (Horn and Cattell, 1967). Knowledge spillovers may then arise from formal and informal interactions within and between firms and might compensate for possible disadvantages of individual ageing. This is particularly true for knowledge-based economies with a higher demand for interactive skills (Autor et al., 2003) that are relatively stable over the life cycle (Skirbekk, 2004). Existing studies on the more aggregated level are far from conclusive though.

At the firm level, studies based on cross-sectional data tend to find a hump-shaped age-productivity profile (Hellerstein et al., 1996; Haltiwanger et al., 1999; Lallemand and Rycx, 2009), whereas studies dealing with the endogeneity of the firm's workforce age by applying panel estimations and instrumental variable techniques suggest either no or even positive effects of older workers on firm productivity (Cardoso et al., 2011; Dostie, 2011; Van Ours and Stoeldraijer, 2011; Göbel and Zwick, 2012). The later findings may hint at the suggested age complementarities within firms and which might be partly compensating for declining mental abilities. In fact, a firm-level study by Backes-Gellner and Veen (2013) argues in this direction and shows that companies involved in creative tasks benefit from an age-diverse workforce.

At the macro level, studies on the link between workforce age structure and regional performance measures are much scarcer although knowledge spillovers between firms have been found to be important drivers of innovation (Feldman and Florida, 1994; Audretsch and

[2]This hump-shaped pattern also seems to hold for general work productivity, see Skirbekk (2004) for a review.

Feldman, 1996). Using more general indicators of economic performance, few country-level studies investigate the link between the workforce age and GDP growth (Lindh and Malmberg, 1999; Prskawetz et al., 2007) or total factor productivity (Feyrer, 2008). Overall, these studies find a hump-shaped pattern even when applying panel and instrumental variable estimators. At the regional level, a hump-shaped pattern has been found by Brunow and Hirte (2006) for GDP growth, by Bönte et al. (2009) for the firm formation rate, and by Frosch and Tivig (2009) for the regional invention rate. Only Bönte et al. (2009) thereby plausibly solve the endogeneity of the regional workforce age by applying instrumental variables. Moreover, none of these studies investigates potential complementarities between different age groups.

This paper fills this research gap by examining the causal link between workforce age structure and patenting activity on the level of local labour markets and by investigating potential complementarities and substitutabilities between different age groups using flexible knowledge production functions. By doing so, the paper makes at least three contributions. First of all, we investigate the link between creative performance and age at the preferred unit of analysis. Previous studies have shown that the link between innovative inputs and outputs appears to be modelled best at a regional level, see Audretsch and Feldman (2004) for a detailed discussion. The relevance of the regional context for the generation of ideas appears to be driven by the spatially limited range of between-firm knowledge externalities which turns the regional level to the preferred unit of measuring the generation of innovations (Peri, 2005). Secondly, analyzing the age-creativity link for German labour market regions is of particular interest since Germany has the second highest median age behind Japan[3] and, more importantly, is characterized by a striking demographic polarization across regions (Gregory and Patuelli, 2013). Thirdly, we address the endogeneity of the regional workforce age by using long lags of the regional population age structure, the share of bohemians and the public sector share for an Instrumental Variables (IV) approach. In addition, we compare both cross-sectional and panel estimations and control for potentially confounding factors such as public and private R&D expenditures, the number of creative professionals, population density and the regional industry mix. We then run various specifications to shed light on the nature of the knowledge production function. For ease of comparison with many existing studies, we first estimate the age-creativity link by using age polynomials in order to derive the age-innovation profile. We then estimate a

[3]See `http://esa.un.org/unpd/wpp/Documentation/pdf/WPP2012_HIGHLIGHTS.pdf`.

Translog production function using the number of young, middle-aged and older workers to gain insights into the complementarities and substitutabilities between these input factors.

Overall, our results suggest a more complex pattern compared to the typically found hump-shaped age-innovation profile from existing studies. In particular, our findings indicate that younger workers boost regional innovations, but that this effect partly hinges on the presence of older workers. Moreover, cross-partial derivatives from Translog production functions suggest that abilities of younger workers and the experience of older workers are complements in the production of knowledge. Despite this positive indirect effect of older workers on the production of knowledge, however, our findings point towards a reduced innovation level if demographic ageing shrinks the size of the younger workforce considerably. Moreover, the difference in the age structure of the least and most innovative German regions explains around a sixth of the gap in innovative performance.

The paper is structured as follows. Section 2.2 discusses the spatial knowledge production function and gives a short literature review on relevant empirical evidence. Section 2.3 introduces the data which is briefly described in Section 2.4. In Section 2.5 we describe the econometric approach before presenting the results in Section 2.6. Section 2.7 concludes.

2.2 Regional Knowledge Production Function

The starting point for our analysis is the knowledge production function which originally has been thought of as operating on the firm level (Griliches, 1979). The knowledge production function describes the relationship between innovative inputs and outputs with R&D investments typically viewed as a main input factor. Whereas empirical studies at the country and industry-level confirm the link between R&D and innovations though (Scherer, 1983; Griliches, 1987; Acs and Audretsch, 1990), the link seems to be much weaker at the firm-level, thus indicating the presence of knowledge spillovers that go beyond the firm (Audretsch and Feldman, 2004). At the same time, such spillovers have been argued to be locally bounded since the transfer of knowledge seems to be linked to face-to-face interactions (Von Hippel, 1994; Manski, 2000). In fact, Peri (2005) shows that only 20% of technological knowledge is learned outside the region. Hence, the natural unit of measuring the generation of innovations appears to be the region, thus giving rise to the regional knowledge production function.

Regional knowledge production functions have been estimated with different measures of innovative outputs and inputs as well as at different spatial units. Jaffe (1989), for example, establishes a positive link between regional research activities by both private corporations and universities and regional patent activity. Using new product innovations as a measure of innovative output, spillovers from academic research and the relevance of corporate spending on R&D have also been confirmed by Acs et al. (1992). Similar findings have been found for Austrian regions by Fischer and Varga (2003). In addition to R&D, human capital has been added as a major input to the regional knowledge production function. In particular, skilled labour has been considered to serve as a main vehicle for knowledge spillovers (Malecki, 1997; Feldman, 1999). Consistent with this notion, Audretsch and Feldman (1996) find that industries with higher shares of skilled labour have a greater tendency to cluster spatially. Knowledge externalities thus seem to be closely linked to the skilled workforce, a notion that is also confirmed by empirical studies on patent activities in the US (Ceh, 2001).

Our approach considers the age of the human capital base to be a major additional input factor of the regional knowledge production function. In particular, we assume that the innovation output in region i is a function of the age of the human capital base and other region-specific factors S_i that have been found to affect the productivity of the regional innovation system. In particular, we consider S_i to capture regional R&D investments by private and public institutions, the skill mix of the regional workforce, a region's industry mix and the scale and density of the local labour market. All of these factors have been shown to affect the regional production of knowledge. For ease of comparison to studies estimating age-invention or age-productivity profiles, we first estimate a simple knowledge production function quadratic in the mean workforce age in region i in addition to these controls and examine the age-innovation profile at the regional level, i.e. we estimate

$$P_i \quad = \quad \alpha_0 + \alpha_1 MAGE_i + \alpha_2 MAGE_i^2 + \beta S_i + u_i \tag{2.1}$$

where $MAGE_i$ corresponds to the mean age of the regional workforce, β is a vector of coefficients for all regressors contained in S_i and u is the usual error component. However, this approach is highly restrictive and does not allow for further insights regarding the relevance of age complementarities.

Hence, we refine the knowledge production function to distinguish between the number of younger workers between 18 and 29 (A_{1i}), middle-aged workers between 30 and 49 (A_{2i}) and older workers above 50 (A_{3i}). The first group comprises young workers that have just completed their education, but who are relatively unexperienced in the labour market. The second group determines workers with increased experience and high productivity levels. Finally, the last group comprises the age group 50 plus for whom studies have shown that cognitive capacities are starting to decline, but who draw from a large stock of experience and skills on team work behaviour and problem solving in difficult situations. For instance, Börsch-Supan and Weiss (2011) conduct an analysis for an assembly plant of a truck manufacturer and find that older workers, though making more errors, are more able to grasp difficult situations and concentrate on the vital tasks. The results are in line with past evidence that suggests fluid abilities (speed of problem-solving and abstract reasoning) to decrease at older ages, whereas crystallized abilities (ability to use skills, knowledge and experience) remain at high functional levels until late in life (Horn and Cattell, 1967). We argue that the skills and experience of older workers may be complementary to young and relatively unexperienced workers, especially in a knowledge-based economy. In fact, other studies have also argued in favour of such age-related skills and complementarities (Schneider, 2008; Göbel and Zwick, 2012; Backes-Gellner and Veen, 2013).

In order to allow for complex patterns of complementarity and substitutability between the age groups, we start with the most flexible functional form, the CES-Translog production function, and test whether the more restrictive Translog, CES and Cobb-Douglas production technologies are suitable approximations of the CES-Translog. As will be discussed in Section 2.6.2, the Translog production function turns out to be a suitable fit for the production of knowledge. The main estimations later on are thus based on estimating the Translog production function with

$$P = A_1^{\alpha_1} A_2^{\alpha_2} A_3^{\alpha_3} e^{(S+\beta_1 \ln A_1^2+\beta_2 \ln A_2^2+\beta_3 \ln A_3^2+\gamma_{12} \ln A_1 \ln A_2+\gamma_{13} \ln A_1 \ln A_3+\gamma_{23} \ln A_2 \ln A_3+u)} \tag{2.2}$$

where the index i has been dropped for simplicity. The Translog production function allows for non-linear relations and interactions between any pair of inputs, i.e. it allows for a broad range of potentially non-constant substitution possibilities.

Note that the additional determinants of innovative performance S in Equation (2.2) are considered to be exogeneous drivers of innovative performance that are not interacted with the three differently aged labour inputs. Although this is restrictive, especially regarding the potential complementarities between education and experience, we decided to impose this restriction in order to keep a manageable amount of parameters and to ease the estimation of the above equation by means of an IV strategy, see Section 2.5 for details.

We estimate the Translog production function by a log-log specification of Equation (2.2) and calculate the marginal products as well as the degree of complementarity or substitutability between the different labour inputs using the estimated parameters. In particular, we calculate the marginal products of each age group as

$$\begin{aligned}
\frac{\partial P}{\partial A_1} &= \frac{P}{A_1}(\alpha_1 + 2\beta_1 \ln A_1 + \gamma_{12} \ln A_2 + \gamma_{13} \ln A_3) := \frac{P}{A_1} Z_1 \\
\frac{\partial P}{\partial A_2} &= \frac{P}{A_2}(\alpha_2 + 2\beta_2 \ln A_2 + \gamma_{12} \ln A_1 + \gamma_{23} \ln A_3) := \frac{P}{A_2} Z_2 \\
\frac{\partial P}{\partial A_3} &= \frac{P}{A_3}(\alpha_3 + 2\beta_3 \ln A_3 + \gamma_{13} \ln A_1 + \gamma_{23} \ln A_2) := \frac{P}{A_3} Z_3
\end{aligned} \tag{2.3}$$

where Z_1, Z_2 and Z_3 represent the elasticities of patent performance with respect to young, middle-aged, and older workers. Note that Z_j with $j = 1, 2, 3$ determines whether a particular age group increases or decreases the regional productivity in terms of knowledge generation.

We then compute the second order derivative for each input factors that corresponds to the change in the previous marginal product with the size of the particular input factor:

$$\begin{aligned}
\sigma_{11} = \frac{\partial^2 P}{\partial A_1^2} &= \frac{P}{A_1^2}(Z_1^2 + 2\beta_1 - Z_1) \\
\sigma_{22} = \frac{\partial^2 P}{\partial A_2^2} &= \frac{P}{A_2^2}(Z_2^2 + 2\beta_2 - Z_2) \\
\sigma_{33} = \frac{\partial^2 P}{\partial A_3^2} &= \frac{P}{A_3^2}(Z_3^2 + 2\beta_3 - Z_3).
\end{aligned} \tag{2.4}$$

Finally, we compute the degree of complementarity or substitutability between the three input factors by estimating the following cross-partial derivatives

$$\begin{aligned}
\sigma_{12} = \frac{\partial^2 P}{\partial A_1 \partial A_2} &= \frac{P}{A_1 A_2}(Z_1 Z_2 + \gamma_{12}) \\
\sigma_{13} = \frac{\partial^2 P}{\partial A_1 \partial A_3} &= \frac{P}{A_1 A_3}(Z_1 Z_3 + \gamma_{13})
\end{aligned} \tag{2.5}$$

$$\sigma_{23} = \frac{\partial^2 P}{\partial A_2 \partial A_3} \quad = \quad \frac{P}{A_2 A_3}(Z_2 Z_3 + \gamma_{23}).$$

with γ_{jk} ($j \neq k$) as the estimated coefficient of the interaction of the two groups of workers in Equation (2.2). The cross-partial derivative gives the change of the marginal product of age group j for a change in the quantity of age group k. The cross-partial derivative thus yields insights into how the expansion of one age group affects the patent performance of another age group. In particular, any pair of inputs A_j and A_k are complements (substitutes) if $\sigma_{jk} > 0$ ($\sigma_{jk} < 0$).[4] We calculate these cross-partial derivatives at the mean of the sample based on the log-log specification of our knowledge production function and use the delta method to derive at standard errors.

2.3 Data

The following data is calculated at the level of local labour markets as defined by Kosfeld and Werner (2012). This classification comprises 141 local labour markets in Germany that have been functionally delineated based on commuting time[5] and do not necessarily follow political boundaries. For each of these 141 regions, we calculate the number of regional innovations as well as demographic and regional indicators on a yearly basis for the time period 1994-2008.

As a measure for innovative outcome in the regional knowledge production function we use regional patent activity. There are several advantages and disadvantages of using patenting data at the regional level.[6] On the one hand, patent applications are a useful indicator of research and invention activities at the local level, as they include information on the regional origin of inventor activities, i.e. place of residence and therefore indirectly the location of the process of knowledge generation. On the other hand, not every invention becomes the subject of a patent application, nor does a patent necessarily become a marketable product or process.

[4] Alternatively, one might calculate the Hicks partial elasticity of complementarity (HEC) (Sato and Koizumi, 1973). The HEC measures the effect on the relative factor price of two input factors that is induced by changes in the relative quantities of these inputs. However, this is a meaningful interpretation only if we assume that the above production function is actually at the core of the profit maximization by firms operating on a competitive market with given output prices. Since patents are not the output sold at the market, we find it implausible to choose such a measure, but stick to the cross-partial derivatives of the production function in order to assess the complementarity of the labour inputs.

[5] Kosfeld and Werner (2012) use a factor analysis based on the commuting time that is reasonable given the size and attractiveness of the region's center (maximally 45 to 60 minutes), see Figure 2.1.

[6] For a detailed discussion see Giese and von Reinhard Stoutz (1998) and Giese (2002).

Moreover, the reasons for a patent application may not only rest on protecting an invention against unjustified use, but may reflect strategic concerns such as securing and extending regional markets, prestige advertisement and the demonstration of innovative capacity to the economic competitors. Despite these disadvantages, empirical evidence by Acs et al. (2002), who provide an exploratory and a regression-based comparison of the innovation counts and patent counts at the lowest possible level of geographical aggregation, suggests that patents provide a fairly reliable measure of innovative activity. Also, the survey provided by Griliches (1998) concludes that patents are a good indicator of differences in inventive activity across different firms.

For this reason, we use patent data that is provided by the European Patent Office (EPO) in order to measure regional innovations between 1994 and 2008 on a yearly basis. The data contains patent applications both at the applicant and inventor level. Whereas the applicant is the holder of the patent right, the inventors are the actual inventors cited in the document. We focus on patent inventors since we are interested in the spatial distribution of the actual inventors rather than the location of the formal holder of the patent, which is often one of the firm's headquarters. Since patents may have been developed by several inventors located in different regions, we apply a fractional counting approach to assign to every region the respective share of the patent. For instance, an inventor who developed a patent in region i with one further individual working abroad would generate 0.5 patents for region i. As a robustness check, we will also use the number of citations of all regional patents as an alternative, more quality-weighted measures of regional innovations.

For the calculation of the age structure of the regional workforce, we make use of the regional file of the Sample of Integrated Labour Market Biographies (SIAB) from the Institute of Employment Research (IAB). This administrative data set is provided by the German Federal Employment Agency and contains a two percent subsample of all workers that are subject to social insurance contributions by their employers, thus excluding civil servants and self-employed individuals. The data includes individual employment histories on a daily basis and contains, among others, information on the age, education and occupation of workers as well as the labour market region of each workplace. We are thus able to compute age and skill characteristics of the regional workforce rather than the regional population. We consider this to be an advantage because regional innovations should be linked to the regional workforce rather than to those living, but not necessarily working in the labour market area, although the distinction should be of no major concern if most commuting takes place within labour market regions. Furthermore, we restrict the analysis to the employed adult workforce. Although knowledge spillovers are not completely restricted to the employed workforce, it is nonetheless unlikely that unemployed workers will participate in the relevant knowledge interactions. The same holds for underage workers who are typically undergoing a vocational training.

For computing the regional workforce characteristics on a yearly basis, we use annual cross sections at the cut-off date June 30th. In particular, we calculate the regional workforce size, the mean age of the regional workforce as well the number and share of workers between 18-29, 30-49 and those above 50 years of age. In addition, we extract further control variables such as the share of workers in certain industries (16 categories) and the share of low-, medium-, and high-skilled workers. Furthermore, following the arguments laid out by Florida (2002), we calculate the number of creative professionals and bohemians[7] of a region since the generation of ideas and innovation seems to largely depend on creative professionals working in the field of education, engineering and science. Bohemians such as artists and publishers, on the other hand, have been argued to create a local milieu that subsequently attracts creative professionals, a link that we will exploit in our IV approach.

As additional control variables, we use information on population density and public research and development (R&D) expenditures (regular and external funding) as provided by the German Statistical Office (Destatis). Moreover, we use rich data collected by the German Stifterverband, that includes private R&D expenditures at a regional level.

2.4 Descriptives

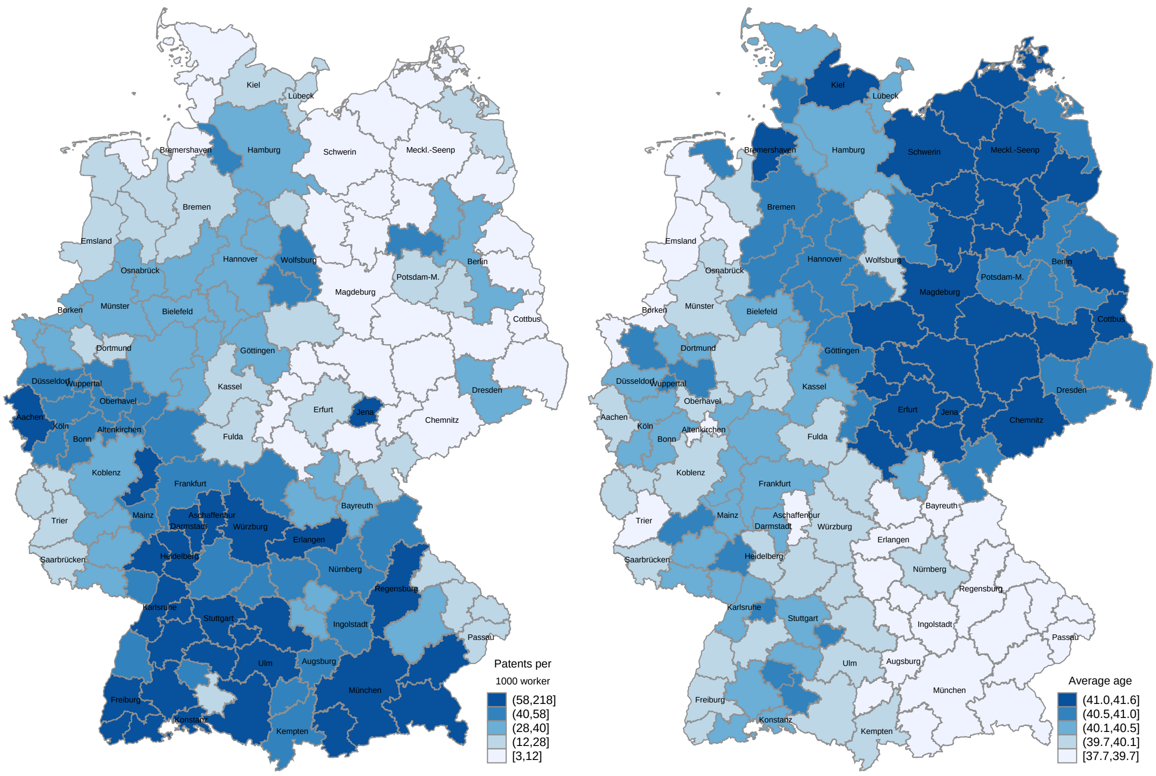
In order to get first insights into the age-innovation link at the regional level, Figure 2.1 maps the mean age of the regional workforce and the regional patent count averaged across the period 1994 to 2008. The regions are classified into quintiles of the respective distributions. Apparently, there are huge cross-sectional differences in the regional patent performance. Whereas the least innovative regions count $3 - 12$ patents per 1,000 workers per year between 1994 to 2008 on average, the most innovative regions score as much as $58 - 218$ patents per 1,000 workers.

[7]For the classification of creative professionals and bohemians, we follow Möller and Tubadji (2009), see Appendix 2.A.1.

Figure 2.1: Workforce age and patent activity by labour market regions (1994-2008)

(a) Number of patents per 1000 worker

(b) Average workforce age

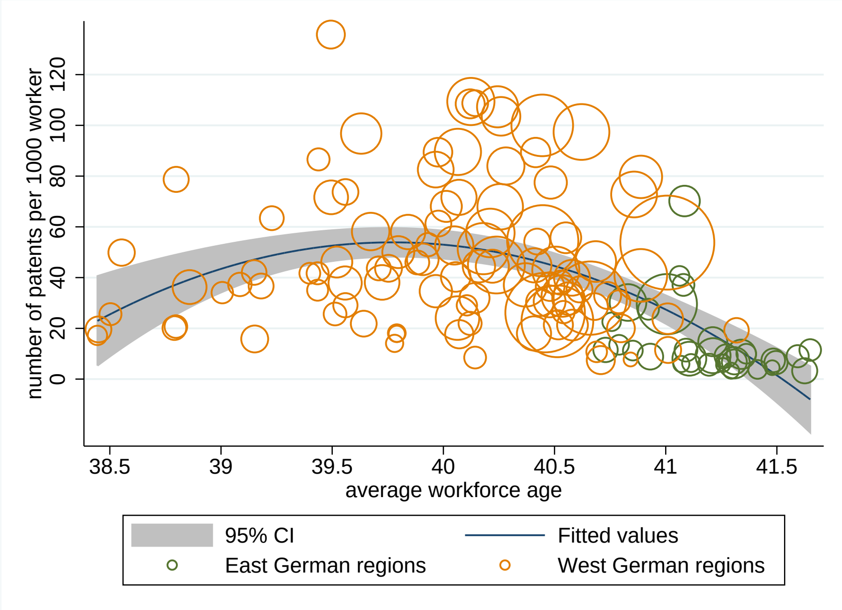


Moreover, these innovation hubs are mainly located in the southern part of West Germany and West German cities such as Duesseldorf, Aachen, Frankfurt, Darmstadt and Heidelberg. In contrast, only a few East German cities such as Jena, Dresden and Berlin seem halfway competitive in the production of knowledge.

Figure 2.1b reveals a large demographic divide between German regions that appears to be highly negatively correlated to regional differences in innovative performance. In particular, the East German workforce is almost two years older on average than the West German workforce, indicating that plant closures and out-migration of young workers after reunification strongly affected the age structure of the East German labour force.[8] Beyond the simple East-West divide, the demographic landscape also seems to reflect an urban-rural divide with many urban areas in West Germany being older than the countryside. Overall, we find substantial regional variation in both the age and innovation dimension that appears to be negatively correlated.

[8]Burda and Hunt (2001) and Hunt (2004) provide empirical evidence for age-selective migration patterns of East-West migration after reunion and discuss the corresponding reasons. A more general approach is taken by Arntz et al. (2014) who define skills as a set of observable characteristics including education and age-related skills of workers and show that Eastern Germany experienced a net loss of such skills during the years between 1995-2008.

Figure 2.2: Scatterplot between average workforce age and patent production, average values for 1994-2008



Notes: The size of the bubbles are proportional to population density. The shaded area represents the 95% confidence interval.

More precisely, a scatterplot for the average workforce age and the average patent count in Figure 2.2 suggests an inversely hump-shaped age-innovation profile when fitting a quadratic relationship with East German regions concentrating at the downward sloping part of the curve.

Of course, the descriptive relationship between workforce age and innovative performance at the regional level may well be driven by other characteristics. Table 2.1 thus contains summary statistics for important control variables by regional patent performance. In particular, we distinguish between the least innovative quintile of all regions and the most innovative quintile and show mean characteristics for these quintiles as well as the respective differences. Whereas the most innovative regions generated, on average, 90.4 patents per 1000 workers, the least innovative regions contributed only 7.7 patents. At the same time, the respective mean age differential between the regional workforces is around one year. More precisely, innovative regions have a higher share of young, but a lower share of middle-aged and especially older workers compared to the least innovative regions. In addition, the innovative regions also seem to be more age heterogeneous as measured by the age dispersion of the regional workforce.

The huge patent gap is, however, not only correlated with the regional workforce age, but also coincides with other well-known drivers of innovation. For instance, the most innovative

Table 2.1: Summary statistics for German labour market regions, 1994-2008

Variable	All regions (1)	Least innovative regions[a] (2)	Most innovative regions[a] (3)	Difference (3)-(2) (4)	Data Source[b]
number of patents	155.77	15.52	359.94	344.41	EPO
number of patents per 1000 worker	40.26	7.74	90.40	82.66	EPO
average workforce age	40.29	41.18	40.05	-1.13	SIAB
workforce age dispersion	10.35	10.15	10.50	0.35	SIAB
share of workers aged 18-29 (in %)	18.05	15.37	18.95	3.58	SIAB
share of workers aged 30-49 (in %)	60.02	60.62	59.25	-1.37	SIAB
share of workers aged 50 plus (in %)	21.93	24.01	21.80	-2.21	SIAB
private RaD expenditures (in 1000 Euro)	243.92	27.35	588.24	560.89	GST
public RaD expenditures (in 1000 Euro)	135.19	47.72	210.90	163.19	DeStatis
share of creative professionals (in %)	5.16	3.72	6.32	2.60	SIAB
share of bohemians (in %)	0.68	0.65	0.76	0.11	SIAB
share of high-skilled workers (in %)	5.83	6.38	6.86	0.48	SIAB
share of medium-skilled workers (in %)	82.70	88.66	78.99	-9.67	SIAB
share of low-skilled workers (in %)	11.47	4.96	14.15	9.19	SIAB
workforce size (in 1000)	3.42	2.10	4.01	1.91	SIAB
population density (population per 100 km^2)	444.98	277.39	516.60	239.20	DeStatis
number of regions	141	28	28	28	

[a] Most (least) innovative regions are defined as regions in the highest (lowest) quintile of the regional innovation (per 1000 worker) distribution.
[b] EPO: European Patent Office, SIAB: Sample of Integrated Labour Market Biographies released by German Federal Employment Agency, DeStatis: Regional database released by Federal Statistical Office, GST: German Stifterverband (Innovation Agency for the German science system)

regions exhibit approximately a twentyfold of private and a fourfold of public R&D expenditures compared to the least innovative regions. Moreover, innovative regions are characterized by a larger workforce, higher population densities and larger shares of creative professionals. Interestingly, innovation hubs show larger shares of both high- and low-skilled workers, but lower shares of medium skilled workers. This might reflect a technology-induced job polarisation as has recently been argued by the task-based literature (Autor et al., 2003).

Adding such controls to our specification of interest, however, will not necessarily ensure the exogeneity of our regressor of interest, the regional age structure. First of all, there may be other time-constant or time-varying omitted variables. Secondly, reversed causality is of major concern since innovative regions might attract young workers. The East-West divide in both demographics and innovative performance might well reflect this inverse link. The following section thus discusses the methodological approach that addresses these concerns and allows for identifying the causal age-innovation link.

2.5 An IV Approach to Estimating the Regional Knowledge Production Function

As briefly discussed in the previous section, only exploiting the cross-sectional variation in our data runs the risk of biases from both time-varying and time-constant omitted variables as well as reversed causality. However, even when exploiting our yearly panel for the period 1994 to 2008, the approach only allows for unobserved time-constant regional heterogeneity so that estimates continue to be biased due to the remaining sources of endogeneity. Moreover, reverse causality is likely to be more severe in the panel dimension. The reason is that the age structure of the workforce that we observe at any given point in time always results from two distinct forces: migration and natural population movements (new cohorts entering and exiting the labour market). To the extent that the regional age structure is inherited from the past due to past economic shocks that are not related to the contemporary innovation activity, but that still affect the contemporary age structure, the reversed causality should be less of a concern. In contrast, changes in the regional age structure over time are more likely to be determined by endogenous forces such as migration. Hence, as suggested by Brunow and Hirte (2006), one approach to mitigate the endogeneity of the age structure is to exploit the cross-sectional variation, since interregional differences in the age structure mainly reflect differences in the age structure of the non-migrant workforce. The advantage of the cross-sectional estimation is that it might be less biased by the endogeneity of (period-wise) migration than a panel approach. On the other hand, any cross-sectional variation may reflect unobserved regional heterogeneities.

Irrespective of whether using a cross-sectional or panel estimation approach, identifying the causal impact of the regional workforce age structure on innovative performance calls for an IV approach in order to mitigate the endogeneity of the regressor of interest. However, instrumenting the age structure in the panel context necessitates a time-varying set of instruments, which is more demanding in the panel than in the cross-sectional context. For all these reasons, we consider a cross-sectional IV regression to be our preferred specification as long as strong and valid instruments can be found. In particular, we consider the following three types of instruments to affect the contemporary age structure of the regional workforce, but to be plausibly exogenous in the innovative performance equation conditional on further controls such as public and private

R&D investments, regional industry mix, workforce size, population density and the skill mix of the regional workforce:

1. The **historical youth-population-ratio** refers to the share of individuals aged 0 to 18 years among the population in region i aged between 0 to 45 years in 1985 as given by the German Statistical Office. This instrument captures the share of individuals that enters the labour market during our observation period and may thus affect the regional workforce age structure. In particular, the higher this youth-population-ratio, the younger should be the regional workforce age. At the same time, we consider this instrument to be unrelated to today's innovative performance since we assume most of these children and teenagers to be born in region i between 1967 and 1985 as families tend to move within local labour markets only (Kulu, 2008). Hence, economic shocks that induced their parents to move to region i before the birth of the first child probably occurred between 1960 to 1980 and are likely to be unrelated to current innovative performance given the structural changes since the late 1970s and 1980s in the aftermath of the oil crises.

2. The **historical share of bohemians** corresponds to the share of individuals that can be considered to be bohemians (e.g. artists, musicians, publishers) among the local workforce as computed from the SIAB data for 1985. The idea behind this instrument stems from the discussion in Florida (2002) who suggests that the localization of the bohemian class is often driven by factors unrelated to economic growth or regional innovative performance, but may than trigger an inward migration of (mostly young) professionals.[9] In line with this research, we assume that the size of the bohemian class 20 years ago rejuvenates the regional workforce 20 years later and is also orthogonal to current innovation output.

3. The **historical public sector share** as measured by the share of public sector workers in 1985 based on the SIAB data, is unlikely to be related to today's innovative performance since the localization of public sector jobs is usually driven by administrative considerations. At the same time, these jobs are typically considered to be particularly family-friendly and thus highly attractive for female workers. In fact, with females increasing their labour force participation throughout the 1970s and 1980s, many women actually entered the public sector. Between 1979 and 2008, for example, the share of women in the public sector in

[9] One example for such a mechanism is Berlin, see Moretti (2012).

the SIAB data increased from 44.8% to 58.7%. We therefore assume public sector hubs to attract young women, thus affecting the regional rate of family formation and, hence, fertility. As a consequence, public sector hubs in 1985 should have a younger workforce twenty years later.

As previously discussed, we will apply these instruments in a cross-sectional estimation of the age-innovation link. However, since the latter two instruments are available for West Germany only, we restrict our main estimations to West Germany. In particular, we estimate two different specifications of the regional production of knowledge:

[A] Knowledge production function quadratic in age. For ease of comparison with much of the literature on age-productivity effects, we begin by estimating the regional patent performance as a quadratic function of the regional workforce age. More precisely, we estimate the following OLS-model for a cross-section of regions where all variables are defined as the average values between 1994 and 2008

$$\ln P_i = \alpha + \gamma_1 MAGE_i + \gamma_2 MAGE_i^2 + \delta \mathbf{S}_i + u_i. \tag{2.6}$$

with $\ln P_i$ as the log of the regional patent count in region i, $MAGE_i$ as the mean age of the regional workforce, and $\mathbf{S}_i$ as a vector of controls including public and private R&D expenditures, the number of creative professionals[10], population density, the structure of the regional industry base measured by the regional employment share of 16 industries and the size of the workforce. When running the estimation for both East and West Germany, we add a dummy for East Germany.

We also run robustness checks for a sample of East and West German regions, but have to restrict the IV set to historical population instruments since our other historical instruments are available for West Germany only. Moreover, we test if our results are robust against using panel estimations by collapsing our yearly data to a panel of five periods, each comprising the average regional value across three years ($t_1 : 1994 - 1996$, $t_2 : 1997 - 1999$, $t_3 : 2000 - 2002$, $t_4 : 2003 - 2005$, $t_5 : 2006 - 2008$). We do so because the yearly patent activity appears to

[10] Alternatively, we used the share of high-skilled workers with a tertiary education, but found only insignificant effects. In fact, the share of creative professionals turned out to be a much more important driver of innovative performance than the level of the formal education.

be strongly varying on a yearly basis whereas changes in the age structure are much more persistent. For this reason, we aggregate three years to one period and allow for a lag between the output and the input measure of one period. We then estimate Equation (2.6) by adding region fixed effects and period dummies. We instrument the endogenous $MAGE_i$ variable by the same set of instruments lagged by three periods, i.e. the mean workforce age between 1994 to 1996, for example, is instrumented by the IV set for 1985-1987. However, the link between these instruments and the change in the regional demographic composition across time is likely to be weaker than in the cross-sectional context.

[B] Translog knowledge production function. Since the use of a quadratic age-innovation link is rather restrictive, we alternatively estimate the Translog production function described in Section 2.2. Compared to Equation (2.6), we use the number of young (18-29), middle-aged (30-49) and older workers (50 plus), its squared terms and interactions. In particular, we estimate the following cross-sectional model again using average values for the period 1994 to 2008:

$$\begin{aligned} \ln P_i &= \alpha + \gamma_1 \ln A_{1,i} + \gamma_2 \ln A_{2,i} + \gamma_2 \ln A_{3,i} + \gamma_3 \ln A_{1,i}^2 + \gamma_4 \ln A_{2,i}^2 + \gamma_5 \ln A_{3,i}^2 \\ &\quad + \gamma_6 (\ln A_1 \times \ln A_2)_i + \gamma_7 (\ln A_1 \times \ln A_3)_i + \gamma_8 (\ln A_2 \times \ln A_3)_i + \delta \mathbf{S}_i + \epsilon_i \end{aligned} \quad (2.7)$$

where A_1, A_2 and A_3 represent our three age groups, and $\mathbf{S}_i$ is a set of controls as defined before except for leaving out workforce size since the size effect is already captured by the sum of our three age groups. With the the three age groups and all its quadratic and interactions terms being endogenous, the IV set as described above does not suffice since we need at least nine instruments for identification. We therefore split up the historical youth-population ratio into five subgroups including the share of 0-3, 3-6, 6-10, 10-15 and 15-18 among the total population in region i aged between 0 and 45 years to allow for more heterogeneity that may affect the share of young, middle-aged and older workers. In fact, the share of 50 plus workers should be driven by the share of the population in middle ages in 1985. For this reason, we further add the share of those aged 18-20, 20-25, 25-30, 30-35, 35-40 and 40-45 in 1985. Of course, for these older workers in 1985, the exogeneity may be more problematic than for the underaged population, but we calculate Hansen j-statistics to get some insights on the validity of the instruments. Also, we add the historical interactions of the share of underaged (0-18), the young (18-30) and the middle-aged (30-45) population that are likely to affect the interacted worker shares in our Translog specification. In addition to the historical population age structure, we complement the IV set with the share of bohemians and the share of public sector jobs as of 1985. We thus have a total of 16 instruments for nine endogenous variables.

2.6 Estimation Results

2.6.1 Age-Innovation Profile

Table 2.2 shows the estimates for Equation (2.6) for West Germany using regional averages for the period 1994 to 2008. Column (1) shows a basic OLS specification with R&D investments, and human capital characteristics only. We then add controls for workforce size, industry shares and agglomeration as measured by population density (Column 2). Columns (3)-(4) use the same set of controls and instrument the mean workforce age and its squared term as described in the previous section. While Column (3) reports the Two Stage Least Squares (2SLS) estimates, the IV regression in Column (4) uses Limited Information Maximum Likelihood (LIML).

First of all, note that our model is able to replicate standard findings of the literature. We find a positive and significant elasticity for private sector R&D expenditures in the range of 0.28-0.38 and an insignificant impact of public R&D investments. This is consistent with other studies on the German regional innovation system. Fritsch and Slavtchev (2007), for example, estimate a random effects panel model and report elasticities of private sector R&D between 0.17 and 0.22, whereas the impact of public R&D is only small. As expected also, the number of creative professionals has a positive sign in Column (1). Although, the share of creative professionals seems to be strongly related to the regional industry mix and the urban density. Once we control for these factors in Column (2), the positive coefficient for creative professionals becomes insignificant.

Regarding the main variables of interest, we find a positive and significant impact of the regional workforce age on innovative performance in the OLS specifications (Columns 1-2). In fact, the coefficients suggest a hump-shaped age-innovation link with a maximum patent activity in regions with a workforce aged 38.9 on average. However, when instrumenting the workforce age in Column (3), the impact turns insignificant, suggesting that the age effect may be driven by

Table 2.2: Cross-sectional estimates for West German regions

Dependent variable: number of patents (log)	OLS		IV	
			2SLS	LIML
	(1)	(2)	(3)	(4)
R&D INPUTS				
private R&D exp. (log, in 100 tsd Euro)	0.38***	0.30***	0.28***	0.28***
	(5.77)	(5.25)	(4.35)	(3.86)
public R&D exp. (log, in 100 tsd Euro)	-0.00	0.01	0.02	0.03
	(-0.05)	(0.27)	(0.66)	(0.67)
HUMAN CAPITAL INPUTS				
average workforce age	17.13***	19.47***	65.58	74.31
	(3.42)	(4.17)	(1.42)	(1.26)
average workforce age (squared)	-0.22***	-0.25***	-0.83	-0.94
	(-3.43)	(-4.18)	(-1.42)	(-1.27)
num. of creative professionals (log)	0.57***	0.22	-0.20	-0.28
	(5.27)	(0.70)	(-0.42)	(-0.48)
REGIONAL INDICATORS				
population density (log)		0.22**	0.47*	0.51*
		(2.26)	(1.92)	(1.70)
workforce size (log, in tsd)		0.37	0.73	0.79
		(1.22)	(1.60)	(1.50)
constant	-339.58***	-394.12***	-1303.39	-1476.46
	(-3.41)	(-4.25)	(-1.42)	(-1.27)
With industry shares?	no	yes	yes	yes
N	108	108	108	108
R-squared	0.911	0.951	0.895	0.873
F	258.5	154.8	57.4	48.8
Hansen (j-statistic)			0.420	0.367
Hansen (p-value)			0.517	0.545

Notes: t-statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered by region. First stage regressions are shown in Appendix 2.A.2.

endogenous forces. The instruments thereby have the expected signs in the first stage, compare Appendix 2.A.2. In particular, the past youth population ratio, the past share of bohemians and the past public sector share have the expected negative and significant signs. Despite the significance of the instruments, the F-Test of excluded instruments is below the rule-of-thumb value of 10. We therefore follow Angrist and Pischke (2009) and re-estimate the model using LIML in Column (4), which is known to be more robust to weak instruments. The estimation parameter do not vary much between 2SLS and LIML though. Also, the Hansen J-statistics suggest that our instruments are valid.

Both IV estimates indicate that the hump-shaped age-innovation profile that we find both descriptively in Section 2.4 and in the OLS estimates in Table 2.2 is driven by endogenous forces. For comparison, we also exploit the panel dimension of our data and compare pooled estimates to a model with region fixed effects both with and without instrumenting the workforce age as described in Section 2.5.[11] According to the estimates in Appendix 2.A.3, the age-innovation link disappears when controlling for unobserved time-constant factors in the fixed effects estimation, irrespective of whether applying an IV strategy or not.[12] In fact, the weak overall explanatory power of the fixed effects model suggests that the regional innovation system and the regional age structure are not changing sufficiently for identifying the effects. Therefore, we think that it is more suitable to exploit the cross-sectional variation across regions that can be instrumented by long lags of the regional population and the other workforce measures.

In addition, we compare our estimates from Table 2.2 to an extended sample that also includes East Germany (see Appendix 2.A.4).[13] In line with our previous estimates, the age-innovation link is similar in the OLS specification, but turns insignificant when using the IV approach. Finally, our estimates also appear robust against using the number of citations of the regional patents rather than the patent count as an alternative outcome measure (see Appendix 2.A.3).

Overall, our estimates based on a knowledge production function with a quadratic specification do not confirm the creativity-diminishing effect of an ageing workforce that has been suggested by many individual- and firm-level studies. Our results rather suggest that the descriptive hump-shaped profile is partly driven by endogenous forces and reverse causality. However, the simple quadratic specification might be too restrictive for a more complex age-related innovation effect. In fact, knowledge interactions between different age groups might take place both within and across firms and could be partially compensating for an age-related decline in cognitive and mental capacities. We therefore investigate the age-innovation link in a more flexible way in the subsequent section.

[11]Note that we loose one of the five periods for the estimation since our output measure is defined as the number of patents in period $t + 1$. This leaves us with four periods and $4 \times 108 = 432$ observations.

[12]The Hansen Test of valid instruments cannot be rejected for the FE-IV model. Also, the corresponding first stage estimates suggest relevant instruments (F-statistics above 10).

[13]As previously discussed, we only have the lagged population structure for instrumenting the mean workforce age when including East German regions in the sample.

Table 2.3: Structural estimates of marginal products, second order and cross-partial derivatives from Equation 2.5 (West-Germany)

Dependent variable in Translog model: number of patents (log)	OLS	IV
	(1)	(2)
predicted patents	196.61***	203.36***
	(8.26)	(7.44)
MARGINAL PRODUCTS (N=108)		
young workers (μ_1)	1.26***	1.43
	(3.39)	(1.60)
middle-aged workers (μ_2)	-0.16	0.10
	(-1.21)	(0.24)
older workers (μ_3)	-0.34	-0.92
	(-1.25)	(-1.21)
SECOND ORDER DERIVATIVES (N=108)		
young workers (σ_{11})	-0.00	0.03
	(-0.13)	(1.55)
middle-aged workers (σ_{22})	0.00	0.00
	(0.78)	(1.02)
older workers (σ_{33})	-0.01	-0.03
	(-0.71)	(-1.07)
CROSS-PARTIAL DERIVATIVES (N=108)		
young and middle-aged workers (σ_{12})	-0.00	-0.02***
	(-1.40)	(-2.87)
young and older workers (σ_{13})	0.01*	0.03**
	(1.91)	(2.06)
middle-aged and older workers (σ_{23})	-0.00	0.00
	(-0.29)	(0.21)

Notes: The table presents estimates at the mean of the sample. t-statistics in parenthesis; calculated based on Delta-method; significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Coefficients reflect absolute increases per 100 additional workers. The corresponding Translog estimates are displayed in Appendix 2.A.6.

2.6.2 Age Complementarities

As a much more flexible alternative to the quadratic specification chosen above, we estimate a Translog knowledge production function[14] as discussed in Sections 2.2 and 2.5 to allow for the possibility that the marginal productivity of one age group actually benefits from the presence of another age group. This might be the case if age-specific strength's turn out to complement each other as could be the case between higher cognitive abilities in younger ages and professional experience and networks of older workers.

[14]In fact, the most flexible production technology would be a CES-Translog functional form. We therefore ran tests whether the more restrictive CES, Translog or Cobb-Douglas production functions are nested in the CES-Translog technology. For Cobb-Douglas and CES, the tests wer rejected, whereas the Translog production technology seems to approximate the more flexible form quite well, see Appendix 2.A.5.

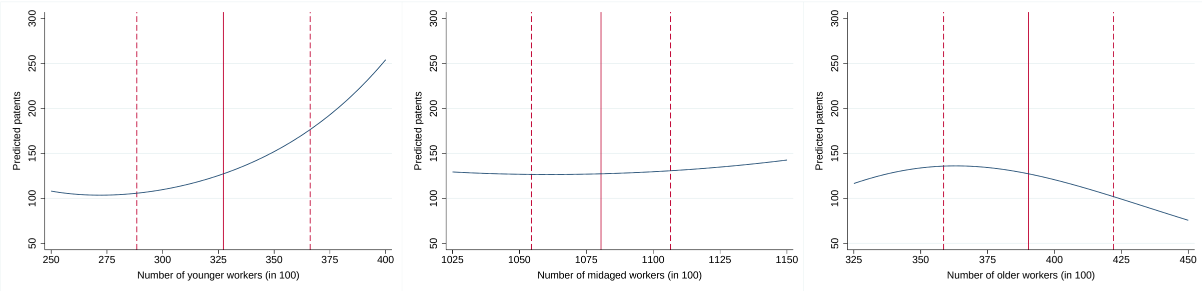
OLS and IV estimates for the Translog knowledge production function from Equation (2.7) are shown in Appendix 2.A.6. We stick to the cross-sectional analysis as our main specification for reasons laid out above. Also, we again restrict the sample to West German regions since we have a stronger set of instruments available for this part of the country.[15] The first stages for these endogenous variables all indicate a strong IV set with F-statistics (of excluded instruments) ranging between 15 and 68.9.[16] Moreover, the Hansen j-statistics strongly suggest that our instruments are valid. We are thus confident that we have both a strong and an exogenous IV set. Based on the OLS and the IV coefficients, we then calculate the marginal products of each age group with respect to the generation of patents (denoted μ_1, μ_2 and μ_3), the second order derivatives (denoted σ_{11}, σ_{22} and σ_{33}) as well as the cross-partial derivatives between the three labour inputs (denoted σ_{12}, σ_{13} and σ_{23}). The estimated structural parameters are summarized in Table 2.3. Note that the marginal products correspond to the effect of extending a particular age group by 100 workers on the number of regional patents while keeping all other input factors constant. The second order derivatives correspond to the change in the marginal product of an input factor with the level of that input. Similarly, the cross-partial derivative, σ_{jk}, refers to the change in the marginal product of age group j due to 100 additional workers of age group k. Note that σ_{jk} equals σ_{kj} in our production function.

Consistent with the previous section, neither OLS nor IV estimates suggest a significant hump-shaped pattern for the age-specific marginal products. Our results rather speak in favour of a more complex pattern although establishing significance for the parameters proves difficult with a sample size of only $N = 108$. In particular, our findings weakly suggest that younger workers boost regional patent activities. According to our estimates, 100 additional workers aged below 30 years increases the patents generated at the regional level by about 1.3 to 1.4, although the effect marginally misses significance in the IV model. The second order derivative σ_{11} is positive which would indicate an increase in this marginal effect with the size of the younger workforce, but, again, the parameter misses significance. For older workers we cannot find a significant negative marginal product on the generation of knowledge. The results are also reflected in

[15] We ran a number of robustness checks, see Appendix 2.A.7. When re-estimating the model based on a sample including East and West German regions, we get quite similar results for the OLS model, but get implausible estimates for the IV model. This likely reflects that we have to drop the historical shares of bohemians and public sector jobs in the IV set when including East German regions, resulting in a much weaker instrument set. When using the number of citations as an alternative outcome measure for a sample of West German regions, however, we get similar results than before.

[16] First stage results are available from the authors upon request.

Figure 2.3: Predicted marginal effects on patent performance by the size of the younger, middle-aged and older workforce.



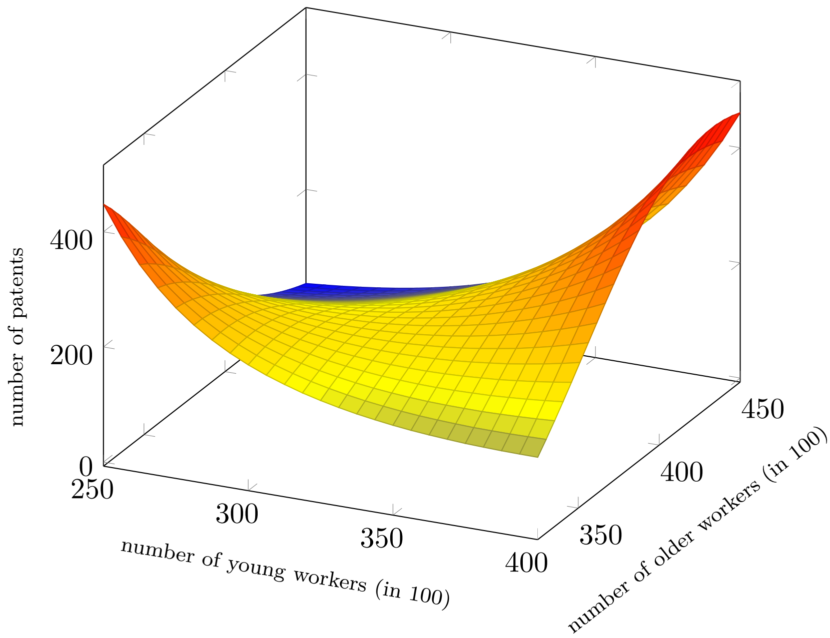
Notes: Prediction based on estimates from IV-Translog specification for West Germany shown in Appendix 2.A.6. The solid red line marks the mean size of the workforce and the dashed lines capture $+/-1$ standard deviation in the share of the respective age group.

Figure 2.3 which plots the age-group-specific marginal effects on regional patenting. The graphs correspond to the predicted effect of each of the three age groups on innovation performance while holding the size of the two other age groups constant at the mean. Not surprisingly given the above estimates, we find a rather flat impact of the size of the middle-aged workforce and a slightly inversely U-shaped pattern for older workers. A sizeable and monotonously increasing effect on patenting only comes from younger workers.

Interestingly, part of this positive marginal effect of a young talent pool results from significant complementarities with older workers. As suggested by σ_{13} based on the IV estimates (see Table 2.3), an increase in the number of older (younger) workers by 100 significantly increases the marginal productivity of younger (older) workers by 0.03 units. The two groups thus boost each others productivity which likely reflects the complementarity of age-specific skills and experiences. The impact of younger workers on regional patenting thus also hinges on the presence of an experienced older workforce. Furthermore, we find some interdependencies between younger and middle-aged workers that indicate a substitutability of these age groups in the invention process. Apparently, these two groups are too similar in skills and experience to benefit from each other.

Simulating patent counts for varying labour inputs. In order to gain more insights into how the detected complementarities between younger and older workers affect the regional production of knowledge, Figure 2.4 plots patent performance as a function of the two labour inputs, holding the size of middle-aged workers constant at the mean. With the current mean region having around 30,000 young and 38,000 older workers, the current mean region is located

Figure 2.4: Simulated patent counts for varying inputs of young and older workers



Notes: The figure plots patent performance as a function of both young and older workers, holding the size of middle-aged workers constant at the mean. Simulations are based on estimates from IV-Translog specification for West Germany shown in Appendix 2.A.6.

in the middle of the graph. There are several things worth noting from this three-dimensional plot. First of all, despite the young worker base having a positive marginal impact on regional innovative performance on average, as shown in Table 2.3, an additional younger worker may even reduce the regional performance if the number of older workers is very low and the negative impact on the marginal productivity of the middle-aged population dominates the positive complementing effect between younger and older workers. With increasing numbers of older workers, however, the marginal effect of younger workers turns positive and even increases sharply for regions with a large amount of elderly. Put differently, regional innovations hinge on the abilities of younger ages, but these abilities need to be complemented by a sufficient amount of experience in order to boost innovations.

As a second observation from the graph, consider the likely movement in the plane that goes along with demographic ageing. In fact, for the time period 1994 to 2008 for which we have data available, the share of middle-aged workers did not change in the average German region. At the same time, the share of workers below the age of 30 dropped by around 6% while the share of 50 plus workers increased by 6% in the average region. Hence, ageing is likely to correspond to a

Table 2.4: Observed, predicted and simulated performance gap between least and most innovative regions with counterfactual age structures

		Least innovative regions	Most innovative regions	Patent gap
		(1)	(2)	(2)-(1)
(A)	observed patents	15.52	359.94	344.41
(B)	predicted patents	52.82***	397.52***	344.7
		(5.68)	(7.14)	
(C)	simulated patents	111.75***	397.52***	285.77
	(assuming age structure of most innovative regions)	(2.83)	(7.14)	
(D)	simulated patents	52.82***	336.77***	283.95
	(assuming age structure of least innovative regions)	(5.68)	(5.58)	

Notes: The table predicts patent performance for the mean characteristics of the least and most innovative regions as shown in Table 2.1. Simulation in Line C (D) predicts patents for the least (most) innovative regions using the age structure of the most (least) innovative region, keeping all other characteristics constant. Calculations are based on IV model in Appendix 2.A.6. t-statistics in parenthesis; calculated based on Delta-method; significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

movement towards the angle in the back of Figure 2.4, a situation that is clearly characterized by low productivity in terms of innovative outcomes. The problem is not so much the increasing number of older workers per se, but the shrinking younger talent pool that may benefit from the experience of the older one.

Explaining the patent gap across German regions. To derive further implications of the estimated parameters for the German innovation divide, we now quantify to what extent differences in the age structure can explain the observed patent gap between the least and most innovative regions. For this purpose, Table 2.4 reports the observed (Line A), predicted (Line B) and simulated (Lines C and D) patent gaps. The simulated values thereby refer to counterfactual predictions for the least (most) innovative regions having the share of young, middle-aged and older workers equal to the most (least) innovative regions, holding other characteristics of the workforce constant. As Table 2.4 shows, the predicted total gap in patent performance amounts to 345 patents which is quite close to the observed gap (compare also Table 2.1). This predicted gap also contains the impact of the difference in the age structure. While the share of middle-aged workers is quite comparable across both types of regions, the most innovative regions have a share of younger workers of around 19% and a share of older workers of around 22% compared to 15% and 24% in the least innovative regions.

When adjusting the age structure in the least innovative regions to equal the structure of the most innovative regions, the predicted patent performance approximately doubles, see Line (C). The remaining gap between the least and most innovative regions that are due to differences in all other characteristics, but hold the age structure constant, however, still amounts to 286 patents which is about 83% of the predicted gap in Line (B). Similarly, adjusting the age structure of the most innovative regions to reflect the shares that we observe in the least innovative regions, reduces the patent gap to 284. Hence, if the least (most) innovative regions had the same age structure than the most (least) innovative German regions, the expected gap in innovative performance would be reduced by about 17%. In terms of economic relevance, this is clearly a sizable impact that again stresses the policy relevance of our results.

2.7 Conclusion

This paper contributes to the debate on demographic ageing in Europe and its potential effects on innovative capacity and regional disparities by evaluating the causal impact of workforce age structure on regional innovations. For the analysis, we estimate a regional knowledge production function quadratic in the mean age of the regional workforce similar to many previous studies on the age-productivity link and contrast this to a much more flexible Translog production technology that allows for complementarities and substitutability between different age groups. We thereby address potential endogeneity of the regional workforce by using long lags of regional demographics, the share of bohemians and the public sector share for an IV approach.

Overall, we do not find a creativity-diminishing effect for ageing labour market regions in Germany in a simple production function with a quadratic specification once taking into account the potential endogeneity of the regional workforce. Hence, the hump-shaped age-innovation profile that is suggested descriptively and by our OLS estimates seems to be driven by endogenous forces. However, when using a more flexible Translog production function, a much more complex pattern is revealed. In particular, our results tentatively indicate that younger workers boost regional innovations, but that this effect partly hinges on the presence of older workers. Moreover, we find a significant complementarity between younger and older workers that may reflect that the cognitive abilities of younger ages and the experience of the older workforce complement each other in the production of knowledge. In contrast, younger and middle-aged workers seem to constitute substitutable input factors in the knowledge production technology. For the least and most innovative German regions, the difference in the age structure is able to explain around 17% of the gap in innovative performance suggesting that our findings are also economically relevant.

For the demographic ageing in Germany and many other western societies, these findings have important policy implications. First of all, our estimates show that up to some point, demographic ageing need not have negative effects on knowledge creation. As long as there are sufficient numbers of younger workers, additional older workers may in fact induce positive effects on innovative performance at first. However, our findings indicate a declining knowledge production in the future if demographic ageing further increases the size of the older workforce at the expense of the younger one. This is because it necessitates a sufficient size of the younger talent pool to benefit from the experience and innovation-enhancing effect of older cohorts. From this we can conclude that certain fears of decreasing creativity in ageing societies appear to be justified. Counteracting effects of complementarities between younger and older workers are both economically relevant and significant, but do not seem to suffice to actually make up for the age-driven disadvantage in the generation of knowledge in case of a continued process of demographic ageing.

Our findings stress the relevance of attracting the young talent from abroad. This is because innovations are likely to increasingly occur only in regions with a relatively high share of younger workers. In fact, from a policy perspective, it may be reasonable to support the development of young hubs as a potential nucleus for innovations. Finally, the complementarities between younger and older workers suggests that policy makers should focus on improving the conditions for an exchange across age groups both within and across firms. Nevertheless, further research is necessary to better understand the factors that favour strong age complementarities and that may counterbalance negative ageing effects. For instance, the role of urban density and agglomeration economies in fostering knowledge exchange between different age cohorts may by exploited. Also, more general production functions may be modelled to identify further cross-elasticities between demographic measures and innovation-enhancing factors such as R&D, skills and the urban context. Finally, investigating the spatial heterogeneity of the effects may provide additional insights into how these demographic trends are shaping the spatial distribution of knowledge creation.

2.A Appendix

2.A.1 Definition of the creative class

SIAB 7508	SIAB-R 7508 (SUF)	Occupational title (German)
Creative Professionals		
621	64	Maschinenbautechniker
622, 623	65	Techniker des Elektrofaches bis Bautechniker
624, 625, 626, 627	66	Vermessungstechniker bis übrige Fertigungstechniker
628	67	Sonstige Techniker
629	68	Industriemeister, Werkmeister
631, 632	69	Biologischtechnische Sonderfachkräfte bis physikalisch-, mathematisch-technische Sonderfachkräfte
633, 634	70	Chemielaboraten bis Photolaboranten
635	71	Technische Zeichner
691, 692	76	Bankfachleute bis Bausparkassenfachleute
751	87	Unternehmer, Geschäftsführer, Geschäftsbe-reichsleiter
752, 753	88	Unternehmensberater, Organisatoren bis Wirtschaftsprüfer, Steuerberater
774	92	Datenverarbeitungsfachleute
Bohemians		
821, 822, 823	99	Publizisten bis Bibliothekare, Archivare, Museumsfachleute
831, 832, 833, 834	100	Musiker bis Dekorationen-, Schildermaler
835, 836, 837, 838	101	Künstlerische und zugeordnete Berufe der Bühnen-, Bild-, Tontechnik bis Artisten, Berufssportler, künstlerische Hilfsberufe

Notes: classification according to (Möller and Tubadji, 2009).

2.A.2 First stage estimates for IV regressions in Table 2.2

Dependent variable:	average workforce age	average workforce age squared
	(1)	(2)
R&D INPUTS		
private RaD exp. (log, in 100 tsd Euro)	-0.01	-1.01
	(-0.18)	(-0.19)
public RaD exp. (log, in 100 tsd Euro)	-0.05	-3.54
	(-1.43)	(-1.41)
HUMAN CAPITAL INPUTS		
num. of creative professionals (in log)	-0.10	-8.52
	(-0.23)	(-0.24)
population density (log)	0.24**	19.69**
	(2.19)	(2.23)
workforce size (log, in tsd)	0.32	25.77
	(0.77)	(0.78)
INSTRUMENTS (YEAR=1985)		
lagged number of bohemians	-0.23**	-18.21**
	(-2.09)	(-2.08)
lagged population ratio	-0.10*	-7.62*
	(-1.85)	(-1.84)
lagged share of workers in public sector	-0.07*	-5.73*
	(-1.95)	(-1.97)
constant	48.90***	2302.77***
	(7.88)	(4.67)
N	108	108
R-squared	0.646	0.646
F	6.7	6.9
F-Test of excluded instruments	4.2	4.2

Notes: t-statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered by region. All models include regional industry shares.

2.A.3 Robustness checks for estimations in Table 2.2 (West Germany)

	Cross-Section		Panel			
Dependent variable:	Number of citations (log)		Number of patents (log)			
	OLS	IV	POLS	POLS-IV	FE	FE-IV
	(1)	(2)	(3)	(4)	(5)	(6)
R&D inputs						
private RaD exp. (log, in 100 tsd Euro)	0.32***	0.30***	0.21***	0.19**	-0.01	-0.02
	(5.35)	(4.32)	(5.74)	(2.54)	(-0.48)	(-0.82)
public RaD exp. (log, in 100 tsd Euro)	0.01	0.03	-0.00	0.02	-0.02	-0.00
	(0.41)	(0.80)	(-0.01)	(0.39)	(-0.62)	(-0.01)
Human capital inputs						
average workforce age	15.55***	59.98	4.61***	-9.42	0.08	-2.74
	(2.92)	(1.38)	(3.37)	(-0.58)	(0.07)	(-1.15)
average workforce age (squared)	-0.20***	-0.75	-0.06***	0.13	-0.00	0.04
	(-2.92)	(-1.38)	(-3.44)	(0.59)	(-0.10)	(1.25)
num. of creative professionals (in log)	-0.15	-0.52	0.33	0.58	-0.23	-0.35
	(-0.44)	(-1.09)	(1.19)	(1.27)	(-0.86)	(-1.03)
Regional indicators						
population density (log)	0.16*	0.37	0.11	-0.16	-0.61	1.52
	(1.70)	(1.54)	(1.23)	(-0.39)	(-0.40)	(0.75)
workforce size (log, in tsd)	0.72**	1.04**	0.40	0.22	0.93	1.19
	(2.14)	(2.28)	(1.50)	(0.51)	(1.32)	(1.34)
constant	-314.05***	-1193.69	-96.73***	165.06	2.87	
	(-2.96)	(-1.39)	(-3.54)	(0.54)	(0.13)	
N	108	108	432	432	432	432
R-squared	0.952	0.912	0.920	0.839	0.407	0.212
F	118.7	57.9	126.9	76.9	12.1	7.2
Hansen (j-statistic)		0.288		0.253		0.009
Hansen (p-value)		0.591		0.615		0.923

Notes: t statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered by region. All IV-models in this table are estimated with 2SLS. All models include regional industry shares.

2.A.4 Cross-sectional estimates for a sample with East and West Germany

Dependent variable: number of patents (log)	OLS		IV	
			2SLS	LIML
	(1)	(2)	(3)	(4)
R&D INPUTS				
private RaD exp. (log, in 100 tsd Euro)	0.44***	0.34***	0.35***	0.36***
	(7.41)	(6.53)	(6.35)	(5.16)
public RaD exp. (log, in 100 tsd Euro)	0.02	0.03	0.03	0.02
	(0.97)	(1.36)	(1.19)	(0.95)
HUMAN CAPITAL INPUTS				
average workforce age	19.90***	21.01***	14.55	10.29
	(3.86)	(4.87)	(0.60)	(0.27)
average workforce age (squared)	-0.25***	-0.27***	-0.18	-0.13
	(-3.88)	(-4.90)	(-0.61)	(-0.28)
num. of creative professionals (in log)	0.44***	0.28	0.31	0.33
	(4.38)	(0.95)	(1.06)	(1.02)
REGIONAL INDICATORS				
dummy for East Germany	-0.45***	-0.11	-0.16	-0.19
	(-3.08)	(-0.57)	(-0.65)	(-0.65)
population density (log)		0.24***	0.22**	0.21**
		(2.92)	(2.57)	(2.03)
workforce size (log, in tsd)		0.16	0.13	0.11
		(0.57)	(0.45)	(0.35)
constant	-394.11***	-424.25***	-296.75	-212.20
	(-3.83)	(-4.98)	(-0.62)	(-0.28)
With industry shares?	no	yes	yes	yes
N	141	141	141	141
R-squared	0.923	0.952	0.951	0.950
F	260.0	164.3	161.9	156.2
Hansen (j-statistic)			1.386	1.288
Hansen (p-value)			0.239	0.256

Notes: t-statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered by region.

2.A.5 Testing different functional forms

	LR χ^2	Prob > χ^2
CES	274.35	0.0000
Cobb-Douglas	15.03	0.0018
Translog	2.81	0.4225
Observations	108	108

2.A.6 Estimates from Translog production functions for Table 2.3

Dependent variable: log number of patents	West Germany		Germany	
	OLS	IV 2SLS	OLS	IV 2SLS
	(1)	(2)	(3)	(4)
R&D investments				
private RaD exp. (log, in 100 tsd Euro)	0.28***	0.31***	0.33***	0.52**
	(4.91)	(4.67)	(6.27)	(2.17)
public RaD exp. (log, in 100 tsd Euro)	0.01	0.02	0.03	0.06
	(0.53)	(0.71)	(1.25)	(0.51)
Human capital inputs				
number of young workers (log)	2.21	39.06	8.04	28.40
	(0.39)	.	(1.23)	(0.24)
number of middle-aged workers (log)	-5.53	-33.38	-11.57	28.35
	(-0.36)	(-1.17)	(-0.70)	(0.08)
number of older workers (log)	4.86	-3.27	4.29	-74.16
	(0.46)	(-0.12)	(0.41)	(-0.34)
number of young workers (log) squared	-1.35	5.51	0.11	1.27
	(-0.86)	(1.44)	(0.08)	(0.10)
number of middle-aged workers (log) squared	3.82	14.38	6.41	-41.81
	(0.62)	(0.99)	(0.90)	(-0.33)
number of older workers (log) squared	-2.84	-15.32	-3.98	-55.62
	(-0.88)	(-1.48)	(-1.24)	(-1.02)
young workers × middle-aged workers	-5.19	-35.21***	-10.34*	-8.65
	(-1.03)	(-4.18)	(-1.84)	(-0.10)
middle-aged workers × older workers	-3.13	6.19	-2.98	102.79
	(-0.37)	(0.25)	(-0.33)	(0.62)
young workers × middle-aged workers	8.67**	24.38***	10.78***	2.36
	(2.39)	(2.83)	(3.01)	(0.03)
num. of creative professionals (in log)	0.00	-0.68	0.30	2.55
	(0.01)	(-1.31)	(1.00)	(0.85)
Regional indicators				
population density (log)	0.28**	0.32*	0.37***	0.52
	(2.52)	(1.82)	(3.50)	(1.04)
dummy for East Germany			-0.14	-0.10
			(-0.67)	(-0.24)
constant	-7.59	9.97	-7.63	25.79
	(-0.76)	(0.47)	(-0.75)	(0.10)
N	108	108	141	141
R-squared	0.958	0.920	0.955	0.708
F		38761.5		48.1
Hansen (j-statistic)		8.620		0.143
Hansen (p-value)		0.196		0.705

Notes: t-statistics in parentheses * p<0.10, ** p<0.05, *** p<0.01. Standard errors are clustered by region. All models include regional industry shares.

2.A.7 Robustness of structural estimates in Table 2.3

	West-Germany (N=108)		Germany (N=141)			
	Citations		Citations		Patent counts	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
predicted patents	361.65***	340.66	278.30***	263.42***	155.21***	154.59***
	(7.08)	(0.00)	(16.89)	(7.58)	(11.17)	(6.58)
MARGINAL PRODUCTS						
young workers (μ_1)	1.72***	1.76	0.98**	1.66	0.83***	1.98
	(2.65)	(0.00)	(2.05)	(0.70)	(3.07)	(1.30)
middle-aged workers (μ_2)	-0.13	0.16	0.22	0.31	0.08	0.22
	(-0.49)	(0.00)	(1.17)	(0.45)	(0.84)	(0.39)
older workers (μ_3)	-0.52	-0.87	-1.17***	-1.66	-0.84***	-1.93*
	(-0.98)	(-0.64)	(-2.91)	(-1.14)	(-3.57)	(-1.93)
SECOND ORDER DERIVATIVES						
young and middle-aged workers (σ_{12})	0.02	0.08	0.01	0.05**	0.00	0.02
	(1.61)	(1.34)	(0.48)	(2.14)	(0.42)	(1.17)
young and older workers (σ_{13})	0.01**	0.02	0.00	0.02***	0.00	0.01
	(2.13)	(0.75)	(1.01)	(3.16)	(0.90)	(1.03)
middle-aged and older workers (σ_{23})	0.00	-0.03	-0.00	0.00	-0.00	0.02
	(0.12)	(-0.60)	(-0.24)	(0.14)	(-0.28)	(0.66)
CROSS-PARTIAL DERIVATIVES						
young and middle-aged workers (σ_{12})	-0.02***	-0.05	-0.01*	-0.03**	-0.00*	-0.01
	(-2.78)	(-1.14)	(-1.74)	(-2.43)	(-1.72)	(-1.25)
young and older workers (σ_{13})	0.03***	0.07	0.02**	0.05**	0.01**	0.02
	(2.68)	(0.97)	(2.54)	(2.30)	(2.12)	(0.58)
middle-aged and older workers (σ_{23})	-0.01	-0.01	-0.00	-0.02***	-0.00	-0.01
	(-1.43)	(-0.28)	(-0.65)	(-3.00)	(-0.46)	(-0.78)

Notes: The table presents estimates at the mean of the sample. t-statistics in parenthesis; calculated based on Delta-method. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. IV-models are estimated with 2SLS. Coefficients reflect absolute increases per 100 additional workers.

3 Can Regional Employment Disparities Explain the Allocation of Human Capital Across Space?[1]

Joint with Melanie Arntz[2] and Florian Lehmer[3]

Abstract: This paper examines the forces driving skill selectivity of regional migration in a context where modelling the migration decision as a wage-maximising process may be insufficient due to persistent employment disparities. The paper thus extends a Borjas type framework on the selectivity of internal migrants to allow migrants to move to regions that best reward their skills in terms of both wages and employment. This framework predicts that high-skilled workers are disproportionately attracted to regions with higher mean wages and employment chances as well as higher regional wage and employment inequalities. Estimates from a labour flow fixed effects model and a GMM estimator show that these predictions hold, but only employment disparities induce a robust and significant skill sorting. The paper thus establishes a missing link why employment disparities may actually be self-reinforcing.

Keywords: gross migration, migration selectivity, wage inequality, employment inequality, regional disparities

JEL-Classification: R23, J31, J61

[1]Published in *Regional Studies*, 2014, Vol. 48, Issue 10, pages 1719-1738.
[2]Universtity of Heidelberg and Centre for European Economic Research (ZEW), Mannheim.
[3]Institute for Employment Research (IAB), Nueremberg.

3.1 Introduction

In contrast to the US, employment rather than wages respond to labour demand shocks in Europe (Blanchard and Katz, 1992; Abraham, 1996; Mertens, 2002). For Germany, Niebuhr et al. (2012) show that the average wage level barely varies across regions and remained nearly unchanged across a period of economic shocks while the unemployment rate turned out to be much more volatile across regions. Similarly, employment disparities have increased in the aftermath of the recent financial crisis with regional unemployment rates for European NUTS3 regions ranging between 0.6 and 33.2 percent in 2010 (Eurostat, 2013). Due to low interregional mobility rates, such disparities tend to be quite persistent (Baddeley et al., 2000; Decressin and Fatás, 1995). Even worse, such disparities may be even self-reinforcing if high employment regions attract a predominantly high-skilled workforce since the local concentration of human capital may trigger a process of cumulative causation due to skill complementarities and local skill externalities (Lucas, 1988; Romer, 1986; Kanbur and Rapoport, 2005; Fratesi and Riggi, 2007).

Yet, the only existing theoretical framework that was proposed by Borjas et al. (1992) links selective migration to wage disparities only. In particular, high-skilled workers ceteris paribus should be attracted to regions that best reward their abilities by paying high wage returns to their skills as reflected in a high wage inequality. Whereas Borjas et al. (1992) and Hunt and Mueller (2004) demonstrate the relevance of a wage-based selection mechanism for internal migration in the US, corresponding evidence in the European context such as Germany has been surprisingly weak (Arntz, 2010; Brücker and Trübswetter, 2007).[2] In a context with strong employment disparities and strong employment responses to economic shocks, this framework may not suffice to explain skill selective migration. In particular, we argue that employment rates are increasing in worker ability and may therefore be seen as a return to investments in skills as also theoretically argued by Helpman et al. (2010).

For this reason, this paper suggests that regional employment disparities may be a missing

[2]The Borjas hypothesis has also been tested for international migration. In particular, Borjas (1991, 1987) provide empirical evidence in favour of the Borjas' Self-Selection Model, assuming migration to be mostly driven by interregional wage differentials. The results have been challenged by Chiswick (1999) and Chiquiar and Hanson (2005) who suggest that migraton costs are likely to be inversely related to earnings, thus leading to different outcomes for immigrant selectivity. For a recent overview and discussion of the literature see Bodvarsson and Van den Berg (2013).

link to explain skill-selective migration. In particular, we extend the Borjas framework to allow for a selection mechanism based on both wage and employment differentials and show that the average skill level of a migration flow is a positive function of the wage and employment inequality in the destination as compared to the origin region. Moreover, unlike the Borjas framework, the model suggests that mean wage and employment differentials also induce a positive skill sorting. As a second contribution, we test these predictions for the average skill level of gross labour flows between 27 German regions. For this purpose, we make use of administrative data which covers nearly 80 percent of the German workforce and determine the skill content of each flow with regard to observable skills. We then regress this skill measure on the mean and the dispersion of the regional wage and employment distribution. Instead of only conditioning on the regional unemployment rate as is done by Pissarides and McMaster (1990), Parikh and Leuvensteijn (2003), Etzo (2011) and others, we thus capture not only the average chances, but also allow regions to differ in how these chances are spread among the local workforce. As a third contribution, we are able to exploit the panel dimension of our data in order to condition on average time constant utility differentials between regions (e.g. amenity differentials) that may otherwise bias the estimation results. In order to control for the endogeneity, we estimate the model with the Difference GMM estimator proposed by Arellano and Bond (1991). The findings confirm the relevance of regional employment disparities for skill-selective migration, while regional wage differentials have no robust and significant impact. Our paper thus fills an important gap in understanding the self-reinforcing nature of interregional employment disparities. Although we focus on interregional migration, the main findings should apply to cross-country migration as well although effects may be weaker due to higher migration costs. Still, our findings suggest that the recent emergence of intra-European migration flows from southern Europe towards high employment countries such as Germany that has been found by Bertoli et al. (2013) is likely skill-biased, thus potentially aggravating the current North-South divide.

The structure of the paper is as follows. Section 3.2 presents an extended theoretical framework for the skill composition of migrants. Section 3.3 introduces the data base, while section 3.4 presents descriptive evidence on the proposed selection mechanism. Section 3.5 describes the estimation strategy and presents the findings which are then subject to additional robustness checks in section 3.6. Section 3.7 concludes.

3.2 Theoretical Framework

Our theoretical framework builds upon Borjas et al. (1992), who formalise the self-selection of interstate migrants and test their model in the US context.[3] While they focus on the selectivity of internal migrants with respect to both observable skills and unobservable abilities, we restrict the analysis to observable skills since we want to understand the spatial allocation of observable skills such as formal education, experience, but also other relevant skills as proxied by occupations which we consider the most important for the skill endowment of a region.

Let Υ be a continuous random variable of a worker's observable skills in the population with mean zero and υ as its realizations. High-skilled workers rank in the upper part of the skill distribution with $\upsilon > 0$ and low-skilled workers rank in the lower part of the skill distribution with $\upsilon < 0$. The distribution is considered to be region-invariant, i.e. we assume that observable skills are perfectly transferable between all regions, an assumption we consider justified for internal migration.[4]

Consider two regions j and k that differ with respect to their (labour) income distribution. As a consequence, migration decisions do not depend on differentials in regional wage distributions alone but also hinge on the probability of receiving this wage, i.e. the probability of being employed as has already been discussed by Todaro (1969) and Fields (1976). For ease of exposition, our theoretical framework abstracts from other utility differentials between regions such as regional amenities or disamenities including regional price differentials as well as from migration costs.[5] An income-maximising individual chooses to live in region j if the expected income in region j exceeds the income in region k, i.e.

$$w_j(\upsilon)e_j(\upsilon) > w_k(\upsilon)e_k(\upsilon) \tag{3.1}$$

with e_j as the individual's chance of being employed in region j on any particular workday and w_j as the wage paid in this region if employed. Note that employment chances e_j are not only

[3]Their framework is linked to the self-selection of workers as described by Roy (1951), and its extension to the self-selection of immigrants as developed by Borjas (1987).

[4]While this assumption may be unproblematic within Western and Eastern Germany, it is less clear whether the assumption can be applied to migration across the former German border. We will thus run some sensitivity analyses in Section 3.6.

[5]Our empirical approach controls for time-constant regional differentials and, thus, takes account of much of these factors.

capturing an initial job finding chance in region j, but should be thought of as measuring the expected probability that an individual is employed on any workday given the region-specific chances of finding, keeping and losing a job.[6] Also note that both wages and employment chances are considered to depend on skills.

In particular, the wage distribution can be decomposed into a part reflecting the mean wage μ_w that is independent of an individual's observable skills and a part that measures skill-specific deviations from this mean wage that depend on the returns to skills paid in region j. Following Borjas et al. (1992), the population wage distribution in region j can then be written as

$$w_j(\upsilon) \quad = \quad \mu_{w_j} + \sigma_{w_j}\upsilon \quad \text{with} \quad \mu_{w_j}, \sigma_{w_j} > 0. \tag{3.2}$$

Hence, an individual's potential wage is determined by their position in the skill distribution and the region-specific returns to these skills σ_{w_j}. Note that even for the least skilled individual, the wage is still positive, thus implying $\mu_{w_j} > \sigma_{w_j}\upsilon$ to hold for every υ. We further assume an analogous decomposition of the population employment distribution, which can be written as

$$e_j(\upsilon) \quad = \quad \mu_{e_j} + \sigma_{e_j}\upsilon \quad \text{with} \quad \mu_{e_j}, \sigma_{e_j} > 0, \tag{3.3}$$

with μ_{e_j} as the average probability of employment on any workday and υ as defined above. Hence, an individual's employment probability is determined by the average employment probability and the region-specific returns to their skill level in terms of employment σ_{e_j}. We thus assume employment chances to be increasing in worker skills, since low-skilled individuals are more likely to be atypically employed (Giesecke, 2009), face a higher probability of becoming unemployed during an adverse regional shock (Mauro and Spilimbergo, 1999) and are more prone to repeated and prolonged unemployment periods (Juhn et al., 1991; Riddell and Song, 2011; Wilke, 2005). In contrast, high-skilled workers are less affected from an economic shock as they are more

[6] Although it might seem unlikely that individual's take account of such abstract measures in their utility maximisation, we argue that individuals gather such information when looking for employment and better jobs across regions. Since individuals mainly move after finding a better job match, migration depends on the chances of getting attractive job offers from a particular region, a probability that can be explicitly modelled in the context of a search model see Damm and Rosholm (2010) for an example. Moreover, search models allow for a simultaneous search across different labour markets and are thus able to explain moves to regions that within a utility framework are suboptimal. Applying a search model would thus be an interesting extension. We still decided to stick to the simpler utility maximizing framework because we model aggregate flows for which the utility maximizing framework gives comparable predictions with regard to wages and employment.

likely to be hoarded during economic downturns and may even become more demanded (Nickell and Bell, 1995; Morrison, 2005). Hence, just as regions with high wage inequality, regions with a high inequality in employment chances penalise low-skilled workers and reward high-skilled workers. The employment dispersion thus measures how the employment chances are spread across the workforce. Since we assume all individuals of the labour force to have positive employment chances, it must also hold that $\mu_{e_j} > \sigma_{e_j}\upsilon$ for all relevant values of υ from the interval $[\upsilon_{min}, \upsilon_{max}]$.

If we apply the two decompositions in Equations (3.2) and (3.3), the income in region j can be written as

$$\begin{aligned} \pi_j(\upsilon) &= (\mu_{w_j} + \sigma_{w_j}\upsilon) \cdot (\mu_{e_j} + \sigma_{e_j}\upsilon) \\ &= \mu_{w_j}\mu_{e_j} + (\mu_{e_j}\sigma_{w_j} + \mu_{w_j}\sigma_{e_j} + \sigma_{w_j}\sigma_{e_j}\upsilon)\upsilon \\ &:= M_j + R_j(\upsilon) \end{aligned} \tag{3.4}$$

where the first term M_j corresponds to the income of an average individual with $\upsilon = 0$ in region j, and the second term $R_j(\upsilon)$ reflects all region-specific returns to skills.[7] This second income component is an increasing function of υ and induces a sorting of individuals into regions that best reward their skills. In particular, the income differential $\Delta\pi$ between region j and k depends on the parameters of the wage and employment distribution in both regions and on individual skills. It can be written as

$$\begin{aligned} \Delta\pi(\upsilon) &= \pi_j(\upsilon) - \pi_k(\upsilon) \\ &= M_j - M_k + R_j(\upsilon) - R_k(\upsilon) \\ &:= \Delta M_{jk} + \Delta R_{jk}(\upsilon). \end{aligned} \tag{3.5}$$

Let S_{kj} denote the average skill level of the migrants moving from region k to j, i.e the skill transfer between both regions. This skill transfer depends on the relative income differentials for workers of different skill levels, i.e. if the income differential increases for high-skilled workers relative to the income differential for low-skilled workers, we would expect the flow of migrants

[7] Note that, analogous to Borjas et al. (1992), we assume relative prices of all skills to be region-invariant so that we do not have to operate with a multifactor model of ability.

to be more skilled on average, i.e.

$$S_{kj} = f(\Delta_\upsilon) = f(\Delta\pi_{|\upsilon>0} - \Delta\pi_{|\upsilon<0}) \quad \text{with} \quad f'(\Delta_\upsilon) > 0. \tag{3.6}$$

Note that we only claim the skill content of the labour flow to be a monotonously increasing function of the income differential, but we do not specify the exact functional link. For the derivation of partial effects, it now suffices to examine how changes in the employment and wage distribution in one region relative to the other affects the income differential Δ_υ between high-skilled and low-skilled workers. Before doing so, let us simplify Δ_υ for the case that high-skilled workers have a skill level of $\upsilon_H = \upsilon_{|\upsilon>0}$ and low-skilled workers have a skill level of $\upsilon_L = \upsilon_{|\upsilon<0}$

$$\begin{aligned}
\Delta_\upsilon &= \Delta M_{jk} + \Delta R_{jk}(\upsilon_H) - (\Delta M_{jk} + \Delta R_{jk}(\upsilon_L)) \\
&= \Delta R_{jk}(\upsilon_H) - \Delta R_{jk}(\upsilon_L) \\
&= (\Delta\mu_{e_{jk}}\sigma_{w_{jk}} + \Delta\mu_{w_{jk}}\sigma_{e_{jk}})(\upsilon_H - \upsilon_L) + \Delta\sigma_{w_{jk}}\sigma_{e_{jk}}\upsilon_H^2 - \Delta\sigma_{w_{jk}}\sigma_{e_{jk}}\upsilon_L^2 \\
&= (\Delta\mu_{e_{jk}}\sigma_{w_{jk}} + \Delta\mu_{w_{jk}}\sigma_{e_{jk}})(\upsilon_H - \upsilon_L) + \Delta\sigma_{w_{jk}}\sigma_{e_{jk}}(\upsilon_H + \upsilon_L)(\upsilon_H - \upsilon_L)
\end{aligned} \tag{3.7}$$

with $\upsilon_H - \upsilon_L > 0$ and $\upsilon_H + \upsilon_L \gtrless 0$.

For the partial effects of increasing employment and wage levels in region j we then get

$$\begin{aligned}
\frac{\partial\Delta_\upsilon}{\partial\mu_{w_j}} &= \sigma_{e_j}(\upsilon_H - \upsilon_L) > 0, \\
\frac{\partial\Delta_\upsilon}{\partial\mu_{e_j}} &= \sigma_{w_j}(\upsilon_H - \upsilon_L) > 0.
\end{aligned} \tag{3.8}$$

According to Equation 3.8, the partial effect of an increasing wage differential (μ_{w_j}) on migration selectivity (Δ_υ) depends on how employment chances are distributed among individuals in the destination region (σ_{e_j}). Phrased differently, the wage-based selection mechanism depends on the skill-specific chances of finding a job in the destination region in the first place. An increasing employment dispersion thereby reflect increasing employment chances for high-skilled workers while at the same time posing a penalty for low-skilled workers from lacking these skills. As a result, the migration flow from k to j should become more skilled on average. The same argument can be applied to an increase in the mean employment chances.

The partial effects for an increasing employment and wage inequality in region j relative to region k can be written as

$$\begin{aligned}\frac{\partial \Delta_\upsilon}{\partial \sigma_{w_j}} &= [\mu_{e_j} + \sigma_{e_j}(\upsilon_H + \upsilon_L)](\upsilon_H - \upsilon_L) > 0 \ , \\ \frac{\partial \Delta_\upsilon}{\partial \sigma_{e_j}} &= [\mu_{w_j} + \sigma_{w_j}(\upsilon_H + \upsilon_L)](\upsilon_H - \upsilon_L) > 0. \end{aligned} \tag{3.9}$$

Note that the sign of these effects is always positive because the values of $(\upsilon_H + \upsilon_L)$ fall in the interval $[\upsilon_{min}, \upsilon_{max}]$ for which it must hold that $\mu_{e_j} > \sigma_{e_j}\upsilon$ and $\mu_{w_j} > \sigma_{w_j}\upsilon$. Hence, an increasing wage inequality rewards high-skilled workers relative to low-skilled workers to the extent that the destination region rewards them relative to low-skilled workers in terms of better employment chances. The strength of this effect, however, depends on the sign of $(\upsilon_H + \upsilon_L)$. One way to interpret this finding is to think of $(\upsilon_H + \upsilon_L)$ as the total skill endowment of the home region. If the region is skilled relative to the national skill distribution, i.e. $(\upsilon_H + \upsilon_L) > 0$, an increasing wage inequality in the destination region will induce a stronger increase in the skill level of the migration flow than if the region had a below average skill endowment. The same arguments apply to the effect of an increasing employment inequality. The skill content of the labour flow from region k to j should consequently become more skilled on average if either wage inequality or employment inequality increases.

Several issues warrant a short discussion. First of all, whereas our theoretical model predicts both wage and employment differentials to influence the selectivity of migration flows, in practice, we expect employment to be more relevant in a German setting with regional wage rigidities. In particular, collective wage agreements in Germany are typically met by industry-specific trade unions and employer associations. Although such agreements are made at the regional level, the agreements obtained in the south-western part of Germany for the metal industry provide an important benchmark for the rest of the economy resulting in regional wage rigidities (Ochel, 2005). In fact, Mertens (2002) finds wages to be rigid across regions compared to the US. According to Schöb and Wildasin (2007) one reason for such regional wage rigidities may be the low level of regional mobility in Europe as compared to the US. Their theoretical model implies that regional labour demand shocks result in strong regional wage disparities if mobility costs are high. This gives rise to wage bargaining regimes that introduce regional wage rigidities in order to mitigate wage risks that are associated with regional labour demand shocks. Given the theoretical and empirical evidence in favor of regional wage rigidities in Germany, we thus expect employment rather than wage differentials to drive skill-selective migration in Germany.

Secondly, our empirical analysis deviates from the theoretical model in that we examine flows in a multi-region context. Still, the mechanisms derived based on the two-region model should be applicable to this broader context. Moreover, it may be helpful to discuss the reasons why we expect flows in opposing directions to exist. This may be the case because there are new cohorts entering the labour market each period among which a certain share is likely to be mismatched to their origin region in terms of their skills. Thus, while the most able individuals will leave their region for regions with higher returns to their skills, low-skilled individuals may prefer the opposite direction in order to minimise the penalty from lacking such skills. Moreover, individuals for whom a particular region once offered the optimal return to their skills need not be optimally matched forever if individuals shift their position in the skills distribution due to training effects or due to the depreciation of skills.

Finally, the model abstracts from a number of potential complications such as regional amenity differentials, price differentials[8] as well as from migration costs. As long as these components are not correlated to the skill level, the key results remain unchanged. However, there are reasons to believe that migration costs decrease in abilities if abilities facilitate the gathering of information and reduce the psychological costs of migration. If this is the case, the key results of the theoretical model remain unchanged only conditional on such costs. Similarly, if individuals differ in how they value certain regional amenities and disamenities depending on their skill level as has been argued by Berry and Glaeser (2005), the key results also remain intact only conditional on regional differentials in amenities and disamenities. Finally, a recent study by Moretti (2013) suggests that real wage differentials may vary across skill-groups, due to skill-specific differences in the cost of living. Our estimation approach, thus, needs to take account of these complicating factors.

[8]In Germany, individual taxes are mainly invariant across regions apart from a minor tax for church affiliation. We thus abstract from taxes in our framework.

3.3 Data

We use the employment register data (BeH) of the German Federal Employment Agency, an administrative data set that contains information on the population working in jobs that are subject to social insurance payments, thus excluding civil servants and self-employed individuals. The data allows to reconstruct individual employment histories including periods of employment and periods of unemployment benefit receipt on a daily basis. For each employment period, the data contains individual and firm-level characteristics including the daily gross wage, the educational attainment as well as the micro-census region of the workplace. We are thus able to identify gross labour flows by comparing workplaces before and after an interregional job transition.

The sample is restricted to the time period between 1995-2004 because the labour flows between Eastern and Western Germany are severely underestimated in the years right after reunification due to the fact that many individuals did not show up in the data before taking up employment in Western Germany. From the mid 1990s on, the observed labour flows correspond to migration patterns that are officially reported by the Federal Statistical Office. This is in line with other studies that find population and labour migration to yield similar results in migration models using German data (Leuvensteijn and Parikh, 2002). We are therefore confident that our labour flows can be considered to proxy quite well for population flows. Furthermore, we focus on men between the age of 16 and 65 because women exhibit a lower labour force attachment than men and move for different reasons since they are often tied to the migration decisions of their spouses.[9]

For all subsequent analyses, we distinguish between 27 aggregated planning districts. These regions lump together 97 German planning districts ('Raumordnungsregionen') that are defined according to commuting ranges and already comprise labour market regions that are relatively self-contained. In order to ensure a sufficient number of job moves between each region for different skill levels, we had to aggregate these planning districts to 27 larger regions. We do so based on an algorithm that minimises the remaining external commuting linkages subject to merging only up to four adjacent regions, thereby ensuring that the regional division yields

[9] We further exclude men attending military or civilian service since they are centrally registered so that the identification of their exact location is not possible, and we neglect apprentices and all employment spells with minor employment since its definition changed in 1999.

relatively equally sized and self-contained labour markets.[10] For each year between 1995 and 2004, we estimate the employment and wage distribution for the 27 regions as well as the size and composition of the 702 gross labour flows between these regions. The following subsections discuss the corresponding details.

3.3.1 Data on Interregional Labour Flows

For the computation of interregional labour flows, we exploit information on almost the entire working population, i.e. we use the full employment register data (BeH). We thereby use yearly cross sections to the cut-off date June 30th and compare the workplace location between two consecutive years. We are thus able to calculate the gross labour flows by identifying the origin and the destination region for all interregional job moves. Note that the identification of an interregional job move necessitates an individual to be employed on June 30th of two consecutive years. Hence, long-term unemployed persons are underrepresented in our data but one should be aware that the sample includes individuals who have been unemployed between these two cut-off days. In total, we observe almost 137 million individuals between 1995 and 2004 of which 3.6 million (2.6%) experience an interregional job move between two consecutive years.

Based on these data, we calculate the average skill level of each gross labour flow, by calculating an alternative skill measure based on ranking individuals in the predicted wage distribution as has been proposed by Borjas et al. (1992).[11] The underlying idea is that wages reflect the marginal product of labour and may thus proxy for abilities and skills. The quality of the corresponding measure, however, depends on the skill characteristics used in the wage regression. Recent studies point towards the relevance of particular types of skills that are used in certain occupations. In particular, Florida et al. (2012) and Bacolod et al. (2009) show that it is primarily analytical skills that receive an urban wage premium. It is therefore important to find proxies for these different skills. More precisely, we estimate individual i's daily gross

[10] Details on the algorithm are available from the authors upon request.

[11] Ideally, we would rank individuals in the income distribution. However, we are not able to estimate the income distribution for the full BeH data because the data is reduced to a cross section that lacks information on the previous employment history. Extending the data to include the full employment history is impossible due to the resulting size of the data.

wage[12] in region j and year t over the time-period 1994-2004 with the following Fixed-Effects (FE) model for all individuals in the sample:

$$log\ w_{ijt} \quad = \quad \beta_0 + \beta_1 REGION_j + \beta_2 YEAR_t + \beta_3 X_{ijt} + c_i + u_{ijt} \tag{3.10}$$

where c_i captures the individual-specific, time-constant unobserved effect and u_{ijt} corresponds to a remaining idiosyncratic error term. The wage is a function of a vector of dummy variables indicating the workplace location (REGION), a vector of dummy variables indicating the year of the observation (YEAR) and a vector of individual- and time-specific observable skill characteristics (X) described above. We therefore include not only age, age squared and educational attainment to proxy for experience and formal education, but also include an individual's occupation (12 categories), industry sector (28 categories) and the establishment size including its squared term. The idea is that the skill mix applied by an individual is well proxied by the occupational status and the industry affiliation. In addition, establishment size might also affect the applied skill mix.[13] In addition, we control for part-time employment because we do not observe hourly wages but only daily wages that may differ between fulltime and part-time employees due to different working hours.

We then predict the wages for all workers in the sample based on the vector of observable skill characteristics (X) only. This way our skill measure does not reflect differences in predicted wages due to region- and year-specific factors. We thus construct a region- and time-invariant skill distribution. We then measure the skill content S_{kjt} of each labour flow by calculating the average predicted log wage for the N movers of each labour flow:

$$S_{kjt} \quad = \quad \frac{1}{N} \sum_{i=1}^{N} \log \hat{w}_{ikjt} \tag{3.11}$$

Note that Equation (3.11) is the empirical operationalisation of the average skill level of migrants following a particular migration flow in Equation (3.6). In order to compare movers and stayers,

[12] Unfortunately, around 15% of all wages are top-coded at the contribution limit of the social security. Therefore, we impute the censored wages with an estimation procedure described by Gartner (2005). This procedure adds a randomly drawn error term to the predicted wage level and, thereby, avoids a strong correlation between the error term and the explanatory variables.

[13] Ideally, we would observe the skills used by each individual. On the other hand, including occupation and industry dummies in addition to establishment size should allow for capturing the skill content quite precisely although we are not able to explicitly pinpoint the types of skills that are rewarded in the labour market as has been done by Bacolod et al. (2009) and Florida et al. (2012).

Table 3.1: Summary statistics for gross labour flows and employees staying in the sending region, 1995-2004

Variable		Mean	SD.	Min	Max	Obs.
Gross labour flows between $k = 1, \ldots, 27$ sending and $j = 1, \ldots, 26$ receiving regions						
Average number of migrants	overall	506	827	5	11955	$K \times J \times T = 7020$
	between		805	15	8700	$K \times J = 702$
	within		193	-1710	5554	$T = 10$
Average predicted daily gross wage[a]	overall	79.4	8.1	49.5	132.3	$K \times J \times T = 7020$
	between		4.9	65.2	94.7	$K \times J = 702$
	within		6.5	50.7	120.5	$T = 10$
Immobile employees in the sending regions $k = 1, \ldots, 27$						
Average number of stayers (in 1000)	overall	494	211	150	1112	$K \times T = 270$
	between		213	170	1050	$K = 27$
	within		24	368	634	$T = 10$
Average predicted daily gross wage[a]	overall	67.6	4.9	56.1	82.1	$K \times T = 270$
	between		3.5	62.4	75.5	$K = 27$
	within		3.5	60.3	75.7	$T = 10$

[a] Average gross daily are calculated as $(\frac{1}{N}\sum_{i=1}^{N} \hat{w_{ikjt}})$.

we also calculate the average predicted log wage for the stayers in the sending region.

Note that we focus on observable skills only although one could also calculate the unobserved skills of migrants by estimating the time-constant unobserved effect c_i in Equation (3.10), as proposed by Borjas et al. (1992). However, one problem with this approach is that unobservable skills and their region-specific returns are not really separable. Since motivation and the like are remunerated differently across regions, c_i differs depending on the regions in which an individual is observed. Put differently, one cannot really construct a region-invariant distribution of unobservable skills. For this reason, we decided to stick to observable skills only.

Table 3.1 reports descriptives on the number of migrants and the average predicted daily gross wage as a proxy for the average skill level for the 7,020 gross labour flows across the ten-year period. On average, 506 migrants with a predicted daily wage of 79.4 Euro follow a particular migration path in any year and are thus a positive selection with regard to observable skills compared to immobile workers whose average predicted daily wage of 67.6 Euro is shown in the bottom panel. Also note that the variation in the migrants average skill level across flows and time is large. This suggests both a substantial spatial re-allocation of human capital at a particular point in time as well as a responsiveness of the skill content of a particular migration path to changing (economic) conditions. Finally, note that the number of migrants is low for some rather distant origin-destination pairs. In order to check whether flows with only few migrants produce outliers that dominate the estimates, we ran sensitivity tests by excluding labour flows with less than 50 migrants (7.96 per cent of all flows).

3.3.2 Regional Wage and Employment Distributions

In order to test the theoretical predictions presented in Section 3.2, we need to estimate the means and standard deviations of the wage and employment distribution for each region and year.

For the construction of the regional wage distribution, we predict the wages of the regional workforce that result from separate region- and year-specific OLS wage regressions analogous to Equation (3.10). By estimating this model separately across years and regions, we allow for varying returns to observable skill characteristics across years and regions. We use the same covariates as in Equation (3.10) since we want to measure the regional differences in the returns to the characteristics that also reflect the skill measure that we use for the labour flows. We then calculate the mean and the standard deviation of the predicted wage distribution for each year and region.[14]

For the regional employment distribution, it is not immediately clear which measure to choose. One might think about using the probability of receiving a job in a particular region, i.e. the job-finding chances. However, for the expected income, the expected probability of being employed in the region on any particular future workday given the region-specific risk of losing a job, being long-term unemployed and finding employment again should be of crucial concern. We do not expect individuals to have rational expectations on this complex probability. Hence, we proxy the relevant employment conditions by the number of days that someone can currently expect to be employed in a particular region. We do so based on a two-per cent random sample of the employment register data since we need full spell information on periods of employment. As with wages, we then construct the predicted employment distribution of the regional workforce. However, since the number of days employed during a year comes with mass points at 0 and 365

[14] The selection of high-skilled individuals into labour markets that best reward their skills as is predicted by the theoretical framework may give rise to an upward bias in the returns to skills as has been shown by Dahl (2002). For this reason, Dahl used bias-corrected returns to skills in an estimation of skill-selective migration. Despite the upward bias in the returns to skills, however, estimation results for the migration model with uncorrected and corrected returns to skills yielded very similar results, which presumably reflects their high positive correlation. We thus refrain from any attempt to correct our estimated returns, especially since transferring the methodology proposed by Dahl is not straightforward in a context where we have repeated polychotomous choices across a ten-year period.

employed days, we need to take account of this unusual distribution by modelling the different cases separately. For this, let $I_{ijt} = 0, 1, 2$ denote an individual-specific indicator function that depends on the number of days d_{ijt} that an individual i is employed during a particular year t in region j:

$$I_{ijt} = \begin{cases} 2 & if \quad d_{ijt} = 365 \\ 1 & if \quad 0 < d_{ijt} < 365 \\ 0 & if \quad d_{ijt} = 0 \end{cases}$$

Individual i's observed number of employed days depends on the probability of being employed all year ($I_{ijt} = 2$), employed between 0 and 365 days ($I_{ijt} = 1$) and being unemployed all year ($I_{ijt} = 0$). According to the law of total probability, the conditional number of days employed in region j at time t can be written as

$$E[d_{ijt}|X_{it}] \quad = \quad P(I_{ijt} = 1|X_{it})E[d_{ijt}|I_{ijt} = 1, X_{it}] + P(I_{ijt} = 2|X_{it})365]. \tag{3.12}$$

The conditional probabilities $P(I_{ijt} = 0|X_{it})$, $P(I_{ijt} = 1|X_{it})$ and $P(I_{ijt} = 2|X_{it})$ are estimated for each region and year by predicting conditional probabilities within a multinomial logit framework. The conditioning set is the same as in Equation (3.10) except for establishment size, which is not available in the two-per cent random sample of the data set. The expected number of days employed conditional on being employed between 0 and 365 days, $E[d_{ijt}|I_{ijt} = 1, X_{it}]$, is estimated running separate region- and year-specific OLS-regressions. Just as with wages, we then calculate the mean and standard deviation of the predicted employment distribution for each region and year. When comparing the official unemployment rate across the ten year period to the share of days not employed that is implied by our employment measure, we found very similar patterns, confirming that our measure captures a meaningful concept.[15] Note that, despite the censoring, there are enough workers employed between 0 and 365 days, so that regions should differ in the standard deviation of employed days, conditional on average employment.

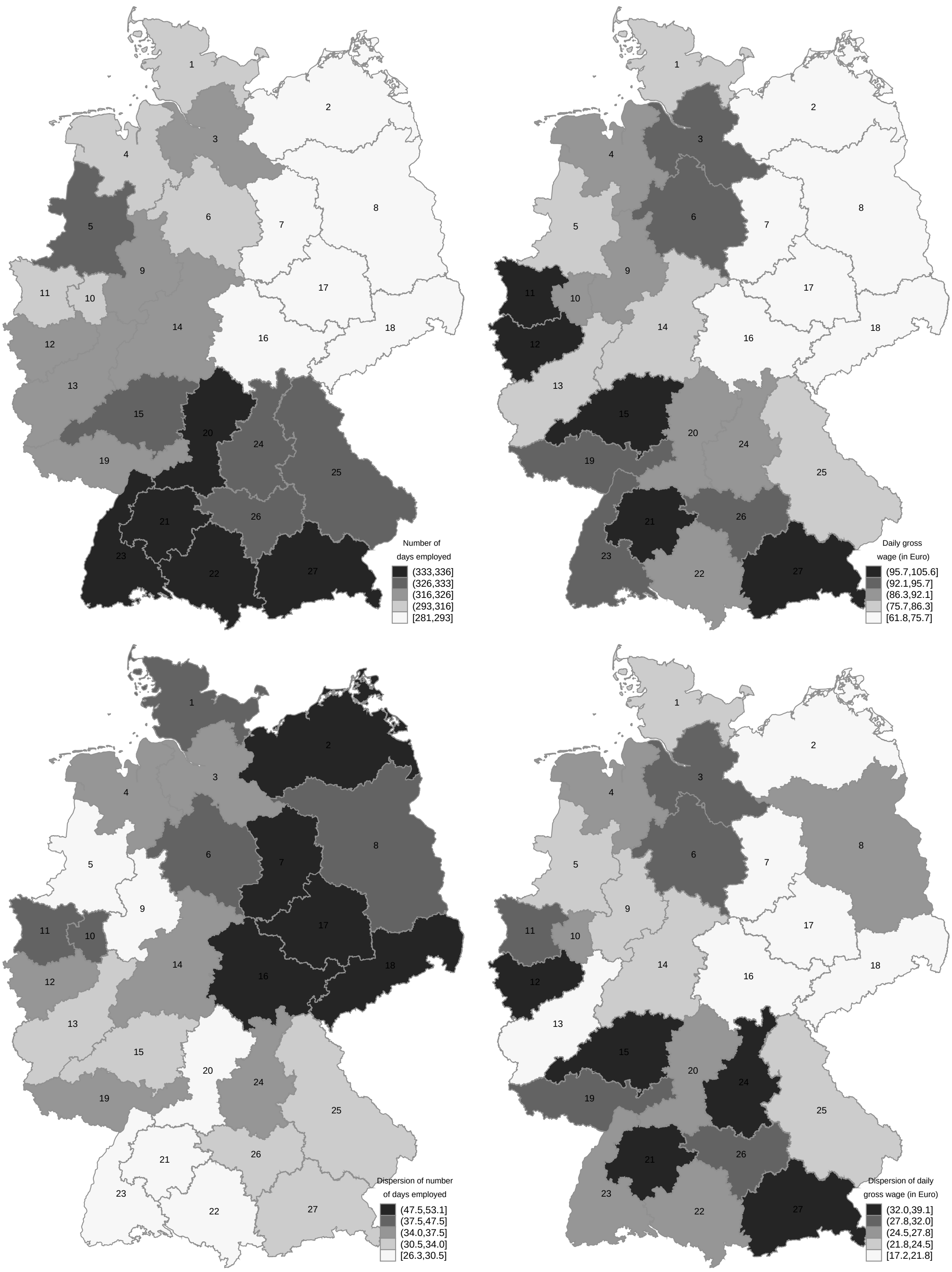
Figure 3.1 shows the average parameters of the employment and wage distribution across the ten year period (absolute changes are shown in Appendix 3.A.1). For better interpretation we use the mean and standard deviation of the exponentiated predicted log wage. We mainly find the

[15] For the definition of unemployment we need to deal with gaps in the employment record, whenever an individual is out of labour force, self-employed, a civil servant or unemployed without any receipt of unemployment transfers. For this, following Fitzenberger and Wilke (2010), we count non-employment periods only as unemployment if there has been at least one initial receipt of unemployment benefits.

expected East-West divide, with average wages and average employment in Western Germany clearly exceeding levels in Eastern Germany. However, we also find some disparities between Southern and Northern Germany, with the latter being in a less favorable labour market situation. Moreover, the absolute wage dispersion in Eastern Germany is below the wage dispersion in Western Germany. Relative to mean wages, though, wage inequality is quite comparable between both parts of the country as has also been suggested by Burda and Hunt (2001) and Gernandt and Pfeiffer (2009) for the time period after 1995. In contrast, the employment dispersion in Eastern Germany strongly exceeds the employment dispersion in Western Germany both in absolute and relative terms. Thus, the risk of being unemployed is not only higher on average in Eastern Germany, but it is also distributed more unequally among the local workforce.

Note that our regional indicators are highly correlated, posing a challenge for the identification of the model. However, since we distinguish 27 regions, there is still a lot of variation on the level of labour flows. Moreover, we exploit the time variation in each of the flows between the 27 regions. Thus, while the East-West divide dominates the picture in a cross-section perspective, our later estimation approach exploits the variation in the skill content of each labour flow across the ten-year period and relates this to changes in interregional disparities. In the subsequent analysis, these refer to the difference between the receiving and the sending region in the standardised wage and employment parameters and are denoted as $\Delta\mu_w$, $\Delta\mu_e$, $\Delta\sigma_w$ and $\Delta\sigma_e$. For the standardization of the regional parameters we subtract the mean and divide by the standard deviation that are calculated across regions and years. An increase of one unit in $\Delta\mu_w$, for instance, thus corresponds to a mean wage one standard deviation higher in the receiving relative to the sending region. Note that we mitigate a simultaneity bias by measuring flows between June 30th of two consecutive years while the wage distribution relates to June 30th prior to observing the flows. Since we need information for an entire year for the employment distribution, the corresponding parameters are estimated for the year prior to observing the destination state in the next year so that there is some but only a limited overlap between the timing of flows and the estimation of the regional employment distribution.

Figure 3.1: Parameters of the regional wage and employment distribution at the level of 27 aggregated planning regions



3.4 Descriptives

If the predictions of the theoretical framework hold, we would expect interregional differences in the mean and dispersion of wages and employment to increase the relative return to migration for high-skilled relative to low-skilled workers. As a consequence, we would expect these returns to increase when ranking labour flows according to their skill content. In order to provide some descriptive evidence on this predicted relationship, we first standardise the observable skill distribution S_{kjt} and calculate the average standardised skill level for each gross labour flow. We then rank all gross labour flows between and within Eastern and Western Germany according to their average standardised observable skill level and create quintiles of this distribution. The lowest quintile among the East-West flows, for instance, corresponds to the 20% of all East-West flows with the lowest skill level, while the fifth quintile captures the 20% of all flows with the highest skill level. We distinguish between flows within and between Eastern and Western Germany to be able to examine whether the relationship is driven by the huge East-West disparities only or whether it holds within each part of the country as well.

Table 3.2 shows the interregional returns to skills measures for the time period 1995-2004 by these quintiles of the labour flows. First of all, note that the average mover even of the least skilled flows in Table 3.2 has an above-average skill level, confirming once again that movers are a positive selection with regard to observable skills. More importantly, we find that within all flow directions most of the parameters respect the expected ranking across the quintiles. Regions offering higher wage and employment returns to high-skilled workers attract labour flows with a higher average skill level. The interregional disparities in employment dispersion, however, are decreasing in skills when the ranking of the mean employment is increasing and vice versa. One explanation may be that regions with high average employment rates tend to have a lower employment dispersion so that $\Delta\mu_p$ and $\Delta\sigma_p$ are strongly negatively correlated with $\rho = -0.51$. We will thus have to run multivariate analyses in order to disentangle the effects. Still, the descriptive evidence tends to confirm that high-skilled migrants move to regions with relatively higher skill premiums in terms of both wages and employment.

Also note that we find differences in the average skill levels across quintiles of different flow directions with flows within Eastern Germany being least skilled. Relating such differences to

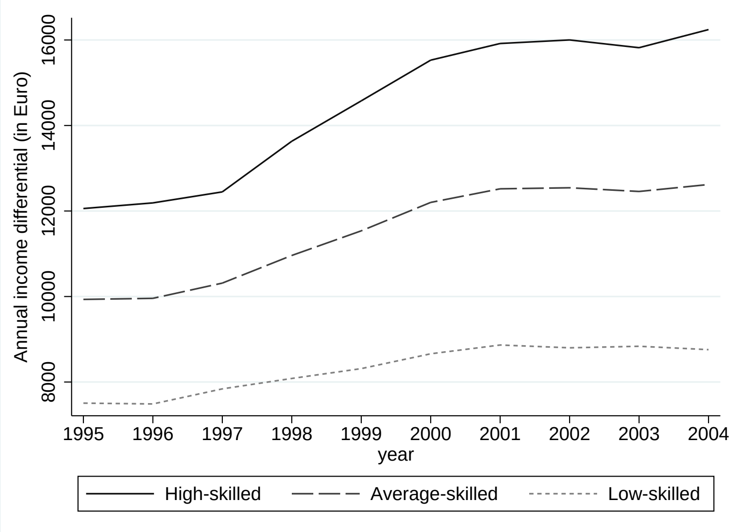
Table 3.2: Average interregional disparities by skill quintile of the labour flows within and between Eastern and Western Germany, 1995-2004

	Observations:			Interregional standardised values:			
Quintile of observable skills distribution	Number of flows	Number of movers (in 1000)	Average standardised skill level	Average wage $\Delta\mu_w$	Average employment $\Delta\mu_e$	Wage dispersion $\Delta\sigma_w$	Employment dispersion $\Delta\sigma_e$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			East-West Flows				
1	251	81	0.11	1.97	2.07	-0.19	-1.67
2	252	101	0.25	1.97	1.92	-0.16	-1.63
3	252	92	0.32	2.13	1.92	0.02	-1.71
4	252	94	0.38	2.23	1.96	0.22	-1.82
5	253	60	0.49	2.37	2.01	0.40	-1.91
			West-East Flows				
1	251	55	0.06	-2.12	-2.13	-0.23	1.72
2	252	69	0.20	-2.09	-2.12	-0.03	1.83
3	252	69	0.27	-2.07	-2.04	0.07	1.84
4	252	60	0.35	-2.18	-1.87	-0.09	1.70
5	253	45	0.48	-2.20	-1.72	-0.01	1.66
			West-West Flows				
1	839	615	0.11	-0.07	-0.08	-0.17	0.10
2	840	607	0.24	-0.07	-0.07	-0.12	0.08
3	840	566	0.30	-0.02	0.01	-0.02	-0.01
4	840	451	0.37	0.05	0.04	0.10	-0.05
5	841	257	0.47	0.11	0.10	0.20	-0.11
			East-East Flows				
1	59	49	-0.10	-0.05	-0.01	-0.12	0.01
2	60	65	0.01	0.03	0.03	0.05	-0.03
3	60	84	0.07	-0.03	-0.05	0.03	0.04
4	60	78	0.13	-0.10	-0.04	-0.21	-0.01
5	61	54	0.21	0.14	0.06	0.24	-0.01
Total	7,020	3,551					

the interregional differences in Figure 3.1, however, may be misleading since the average skill level should also be affected by other factors than regional differentials in wages and employment such as, for example, amenity differentials. For this reason, keep in mind that our aim is not to fully explain the observed skill composition, but to test whether changes in the skill composition of flows are related to changes in the interregional differences in employment and wages as theoretically predicted.

Hence, a better descriptive test is to look at how changes in employment and wage differentials across time are related to changes in the skill composition for any particular labour flow. Of course, such an analysis is not feasible for the 702 available flows. As an example, we thus focus on the flow between Eastern and Western Germany, which is of particular interest given the strong interregional differences that still persist after reunification. Figure 3.2 shows the

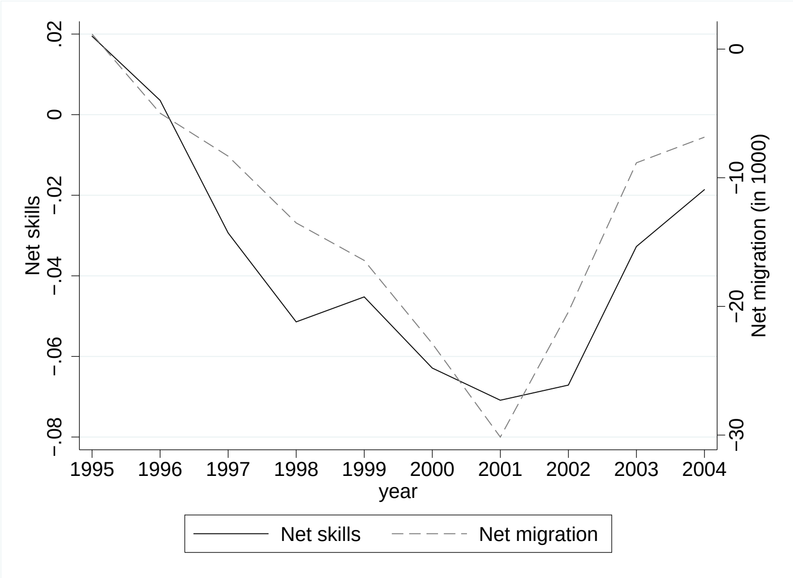
Figure 3.2: (Labour) income differential $\Delta\pi(\upsilon) = \pi^{west}(\upsilon) - \pi^{east}(\upsilon)$ for an average-skilled ($\upsilon = 0$), high-skilled ($\upsilon = 1$) and low-skilled ($\upsilon = -1$) individual, 1995-2004



corresponding (labour) income differential π calculated based on Equation (4) for an individual with average ($\upsilon = 0$), above-average ($\upsilon = 1$) and below-average ($\upsilon = -1$) skills. In 1995, an average individual earns 10,000 Euro per year more in Western than in Eastern Germany. This income differential increases up to 16,000 Euro in 2001 reflecting deteriorating average employment chances in Eastern relative to Western Germany while wage differentials remain rather constant. In 2001, the increase in the income differential came to a halt before stagnating from then on. In light of this development, standard migration models on net migration suggest an increasing net loss of migrants in Eastern Germany until 2001.

The spread between the upper and the lower line in Figure 3.2 reflects the skill premium for high-skilled ($\upsilon = 1$) relative to low-skilled individuals ($\upsilon = -1$) which according to the theoretical framework should drive the skill composition between both parts of the country (see Equation 3.6). The increasing gap between the income differentials of high-skilled and low-skilled workers until 2001 reflects the increasing skill premium in Western relative to Eastern Germany that is mainly related to the deteriorating average employment chances in Eastern relative to Western Germany and, in addition, to an increasing average wage and wage dispersion in Western as compared to Eastern Germany (see also Appendix 3.A.1). Note, that the increasing employment dispersion in Eastern compared to Western Germany obviously didn't close the spread between high-skilled and low-skilled workers (compare Figure 3.2). Hence, according to theory, we should observe an increasing net transfer of observable skills from Eastern to Western

Figure 3.3: Net migration and net skill transfer in Eastern Germany, 1995-2004



Germany similar to net migration.

Indeed, Figure 3.3 shows that the net loss of migrants as well as the net loss of skills in Eastern Germany increased until 2001 and thus confirms the above derived predictions. In particular, in 2001 about 30,000 more migrants moved from Eastern to Western Germany than vice versa. Moreover, the average East-West migrant became more skilled relative to the average West-East migrant, thus aggravating the already existing brain drain with regard to observable skills. While this development is consistent with the development of income differentials until 2001, the somewhat decreasing net loss in migrants and skills after 2001 cannot be fully explained by the income differentials after 2001.

The descriptive findings are in line with the few existing studies by Burda and Hunt (2001), Hunt (2006) and Brücker and Trübswetter (2007) that consider the selectivity of East-West migration flows. In particular, Burda and Hunt (2001) and Hunt (2006) study the years between reunification and the millennium and find that East-West migrants tend to be young and better educated compared to stayers. The authors however do not explicitly study the forces driving the skill composition of East-West migration flows. In contrast, Brücker and Trübswetter (2007) show for the time period 1994-1997 that East-West migrants also constitute a positive selection based on unobservable characteristics.

3.5 Empirical Analysis

In order to identify the determinants of the average skill level of gross migration flows, we exploit the variation in the average observable skill level across 7,020 gross labour flows that we observe during the time period between 1995-2004. The panel is balanced, that is, for all 702 region pairs we have 10 year observations. Since unobserved effects in the error term such as amenity and price differentials are likely to be correlated with regional wages and employment, a simple Ordinary Least Squares (OLS) regression should be biased. Therefore, we estimate the following labour-flow fixed effects (FE) panel model:

$$S_{kjt} = \beta_0 + \beta_1 \Delta\mu_{w_t} + \beta_2 \Delta\mu_{e_t} + \beta_3 \Delta\sigma_{w_t} + \beta_4 \Delta\sigma_{e_t} + c_{kj} + u_{kjt} \quad (3.13)$$

where S_{kjt} refers to the average observable skill of a migrant moving from region k to j in a particular year $t = 1, ..., 10$ with $k \neq j$ as shown in Equation (3.11). The right-hand side of Equation (3.13) contains the interregional differences in the returns to skills, namely the differences in wages and employment between the destination region j and the region of origin k as defined in Section 3.3.2. The composite error consists of the flow fixed effect c_{kj} as well as an idiosyncratic error term u_{kjt}. Hence, the above model controls for time-constant, flow-specific unobserved effects such as time-invariant amenity or costs of living differentials and explicitly allows the explanatory variables to be correlated with such characteristics. We assume remaining biases from time-varying amenity and price differentials to be of minor relevance given the rather short observation-period of ten years.

The results for the skill selectivity are presented in Table 3.3.[16] Models (1) and (2) show the effects of regional wage and employment differentials on the average skill level of a flow measured in log wage points (see Equation 3.11). While Column (1) shows the pooled OLS results that assume time-constant flow-specific factors to be uncorrelated with the covariates, Column (2) shows the above discussed labour-flow fixed effects model with 702 flow-specific

[16] We also estimated Equation (3.11) using absolute migration as a dependent variable in order to test whether our specification replicates migration patterns that have been found in the literature. As expected, we found that increasing mean wages and mean employment chances in the receiving relative to the sending region raises gross migration levels. Moreover, consistent with similar studies on internal migration, employment differentials have a stronger impact than wage differentials (McCormick, 1997; Ederveen et al., 2007; Puhani, 2001). For studies on the particular German case see Parikh and Leuvensteijn (2003) and Decressin (1994).

fixed effects (Equation 3.13). In the pooled OLS model, we add the distance between regions in logs to control for major differences across flows in the costs that are related with moving from k to j. In Column (2), such effects are absorbed in the fixed effects. Some of the coefficients in Column (1) show the expected signs. In particular, distance has a positive effect, indicating that low-skilled workers are relatively more distracted from migrating, compared to high-skilled workers, thus increasing the average skill level of a migration flow. The result suggests that migration costs decrease in abilities since abilities facilitate the gathering of information and reduce the psychological costs of migration (see e.g. Chiswick 1999). The coefficients for mean employment, wage dispersion and employment dispersion also show positive values, although not significant in the case of employment dispersion. Most surprisingly, the regional mean wage differential is negative and significant. This result does not correspond to the theoretical predictions. Instead, adding labour-flow fixed effects in Column (2) yields different and much more plausible outcomes and thus demonstrates that the pooled OLS model is apparently biased by time-constant interregional differentials. For most previous studies that are based on exploiting cross-sectional variation only, this puts doubt on the reliability of the findings and suggests the need for better taking account of unobserved interregional disparities.

In particular, Column (2) shows positive and significant effects of regional differentials in employment. To be more precise, an increase in the mean number of employed days by one standard deviation in the receiving relative to the sending region increases the skill level of the average migrant by 0.027 log wage points. For an average gross wage income of 79.4 Euro per day, this effect would be equivalent to an increase of the daily gross wage of an average migrant by 2.17 Euro. Similarly, regions with a one standard deviation higher employment dispersion attract migrants that are more skilled by 0.023 log points on average. In contrast, coefficients related to the wage level and wage dispersion are small and insignificant. The estimations thus show that in a German setting with regional wage rigidities the skill composition of a labour flow is mainly driven by regional differences in the employment rather than the wage distribution.

Moreover, when comparing these estimates to the results of the standard wage-based selection model as proposed by Borjas and others, see Column (3), their impact remains insignificant. In order to compare the performance of the wage-based and income-based selection model, Figure 3.4 compares the predicted net skill transfer from Eastern to Western Germany based on Models (2) and (3) with the observed net skill loss in Eastern Germany (shown in Figure 3.3). The

Table 3.3: Average skill level of gross labour flows, 1995-2004

Dependent variable: average observable skill of a migrant (S_{kjt})			
	pooled OLS (1)	FE (2)	FE (3)
Δ Mean wage ($\Delta\mu_w$)	-0.008***	0.008	-0.011
	(-3.29)	(0.37)	(-0.56)
Δ Mean empl. ($\Delta\mu_e$)	0.019***	0.027***	
	(6.49)	(6.18)	
Δ Wage dispersion ($\Delta\sigma_w$)	0.012***	0.011	0.004
	(8.45)	(1.50)	(0.57)
Δ Empl. dispersion ($\Delta\sigma_e$)	0.011	0.023***	
	(1.51)	(3.18)	
Distance (log)	0.035***		
	(13.06)		
Constant	4.020***	4.217***	4.217***
	(255.48)	(1192.10)	(1173.54)
N	7020	7020	7020
F	154	168	196
R^2	0.254	0.339	0.332
$Adj.R^2$	0.252	0.338	0.331
$RMSE$	0.112	0.085	0.085

Notes: robust t-statistics in parentheses; Significance levels: * 10%, ** 5%, *** 1%; All models include year dummies. FE denotes labour-flow fixed effects models.

predictions of the income-based selection model perform much better in explaining the observed net skill loss in Eastern Germany compared to the purely wage-based model. The test underlines the importance of extending the framework on skill selective migration to allow for regional employment disparities.

Of course, looking at the impact of regional wage and employment differentials on the skill composition does not tell us much about which migrants drive the results. As an example, we could achieve the same theoretically proposed effects with high-skilled individuals being attracted to regions with a high return in wages or employment, by low-skilled individuals being distracted from these regions, or a combination of both. From the predictions in Section 3.2, we would expect individuals to be increasingly attracted to regions with higher mean wages and mean employment the higher their skill level. In addition, we would expect individuals with below-average skills to be distracted from regions with a high wage and employment dispersion, while individuals with above-average skills should be attracted to these regions.

Figure 3.4: Prediction of East-West migration selectivity based on the wage-based selection model (3) versus the extended income-based selection model (3) in Table 3.3, 1995-2004

In order to test these predictions, we determine the quintile of the observable skill distribution for each worker in our sample and compute respective migration rates for each quintile of the 702 flows across the ten-year period, i.e. we calculate the absolute number of migrants from k to j by quintile of the skill distribution and divide these numbers by the size of the labour force in k of the corresponding skill quintile.[17] The first quintile therefore contains the rate by which individuals ranking in the lowest quintile of the skill distribution follow a particular migration path.

The results are shown in Table 3.4 and can be interpreted as follows: an increase in the mean employment differential by one standard deviation increases the migration rate of the first quintile (20 per cent less skilled migrants) by 18.4 per cent. As expected, the effects of mean wage and employment differentials tend to be positive for all quintiles, although an increasing and significant pattern across quintiles can be found for the mean employment differential only. Thus, the related positive selection in Table 3.3 seems to be the result of migratory responses along the entire skill distribution. In contrast, mean wage differentials significantly affect migration choices for the best skilled individuals only. Since this group is more likely to earn a wage independent

[17]Since we use the log migration rate for our estimations, we add one to the numerator in order to avoid missing for year-flow observations with zero migrants.

Table 3.4: Log migration rates by quintile of the skill distribution, labour flow fixed effects estimation 1995-2004

Dependent variable: log migration rate by quintile of the skill distribution					
	1. Quintile	2. Quintile	3. Quintile	4. Quintile	5. Quintile
Δ Mean wage ($\Delta\mu_w$)	0.130	0.055	0.079	0.100*	0.210***
	(1.50)	(0.78)	(1.23)	(1.66)	(3.30)
Δ Mean empl. ($\Delta\mu_e$)	0.165***	0.216***	0.272***	0.312***	0.321***
	(9.85)	(14.15)	(18.68)	(22.01)	(20.59)
Δ Wage dispersion ($\Delta\sigma_w$)	-0.021	-0.052*	0.024	0.068***	0.033
	(-0.64)	(-1.91)	(0.97)	(2.99)	(1.39)
Δ Empl. dispersion ($\Delta\sigma_e$)	0.096***	0.132***	0.158***	0.157***	0.205***
	(2.84)	(5.07)	(6.33)	(7.29)	(9.87)
Constant	-8.360***	-7.814***	-7.544***	-7.312***	-7.420***
	(-553.65)	(-679.28)	(-726.19)	(-774.30)	(-726.33)
N	7020	7020	7020	7020	7020
F	51	72	118	155	136
R^2	0.105	0.151	0.238	0.321	0.345
$Adj.R^2$	0.104	0.149	0.237	0.320	0.344
$RMSE$	0.360	0.284	0.250	0.215	0.221

Notes: robust t-statistics in parentheses; Significance levels: * 10%, ** 5%, *** 1%; All models include dummies for the origin and destination region.

of any centrally bargained tariff, wage rigidities may be less pronounced for this group. In that case, mean wage differentials are likely to reflect regional differences in income opportunities for the group of high-skilled individuals only, and so it appears plausible that we find significant effects for the fifth quintile only.

Regarding the effect of the wage dispersion, we find no consistent pattern across all five quintiles, but a negative significant effect for the second and a positive significant impact for the fourth quintile that are both in line with the theoretical predictions. In contrast, the employment dispersion has a significantly positive and increasing impact on migration along the entire skill distribution. The related positive selection effect in Table 3.3 thus results from these increasing patterns rather than individuals with below-average skills avoiding regions with a higher employment inequality. While this may seem at odds with the theoretical predictions, we think that this observation is the result of two empirical facts.

First of all, the empirical distribution of the number of employed days turned out to be bi-modal with the majority of individuals being almost always employed. Thus, although the risk of unemployment is spread unequally across individuals, only a rather small share of individuals appears to be affected by a positive unemployment risk. If these individuals were the least skilled as postulated in the theoretical framework, however, we would expect a negative sign for the lowest quintile only and insignificant effects for the remaining quintiles. As an alternative explanation, we think that unobservable skills play a major role. Due to the relevance of unobservable skills in determining an individual's employment chances, even the individual with the highest observable skill level has a non-zero chance of unemployment, while even the individual with the lowest observable skill level may be continuously employed. Therefore, an increasing employment inequality may be beneficial for the majority of individuals with favourable unobservable characteristics within each skill quintile, yet the share of those for whom an increasing employment inequality increases the risk of unemployment decreases across the quintiles of the observable skill distribution. As a result, we obtain continuously increasing pattern of positive coefficients across quintiles in Table 3.4. This pattern thus reveals that the theoretical framework needs to be further extended to selective migration with regard to unobservable skills in order to capture its full complexity.

3.6 Robustness

The main problem with the previous labour flow fixed effect model is that it assumes the independent variables to be uncorrelated with past and possibly current realisations of the error term. However, migration could be both the cause and consequence of regional wages and employment. In fact, several studies explicitly examine the effect of selective migration on the development of regional wages and employment (Fratesi and Riggi, 2007; Burda and Wyplosz, 1992). Although the timing of the dependent and independent variables rule out a direct simultaneity between interregional differences and the skill composition of labour flows, the described dynamics may still bias the previous labour-flow fixed effects estimations.

In order to deal with endogenous regressors and reverse causality, we estimate the average skill level of the gross labour flows with the first-difference General Methods of Moments (GMM) estimator as proposed by Arellano and Bond (1991). The estimator is designed for panel data, where the number of time periods, T, is small and the number of observations, N, is large. The underlying idea is to instrument the endogenous variables in the differenced equation using the lagged versions of the endogenous variables. As Arellano and Bond (1991) note, lagged variables dated t-2 and earlier can potentially be orthogonal to the error and therefore act as valid instruments.[18]

In addition, we not only test this difference GMM for the average skill level of our gross flows, but also run this test for the average skill level relative to the skill level of the sending region. Such a test would be unnecessary if, as assumed by the theoretical framework, there was a skill distribution such that the position of an individual within this distribution is region-invariant. However, this need not be the case, especially between Eastern and Western Germany, where educational systems were separate until 1990. In fact, we do find differences in the average skill level of the regional population, especially between Eastern and Western Germany. While some of these differences may be the result of previous skill-selective migration, such differences may also indicate that the skill distribution is not region-invariant.

Table 3.5 reports the Difference GMM results for both the average skill level and the average relative skill level of a migrant. Thus, in the latter case, migrants may now be high-skilled relative to the source population despite being low-skilled compared to other flows. Still, we would expect the sorting of skills across space to follow the same predicted pattern as before. In all models, we use all available lags of the dependent variable and of the four returns to skills variables dated t-6 and earlier as instruments for the transformed equation.[19] In addition, Models (2) and (4) add lags of the independent variables in order to allow for the possibility that there is a time lag between regional differentials and migratory responses. According to Table 3.5, the orthogonality restrictions of the instruments and the estimated residuals are accepted in all models by the Sargan and Hansen Test. As a test for autocorrelation, we conduct the Arellano-Bond Test on the residuals in differences. The $AR(2)$-Test rejects the hypothesis of autocorrelation of second order which argues against a dynamic model with lags of the dependent variable. In fact, including lags of the dependent variable turned out to be insignificant in all models.

Model (1) shows that results remain quite robust when taking into account both the potential endogeneity of the regressors, therefore confirming the results in Table 3.3. The size of the coefficient for the mean employment differential is, however, slightly higher compared to the basic flow fixed effects model in Table 3.3. The same pattern holds when using the relative skill

[18] The model is also estimated with the System GMM estimator proposed by Blundell and Bond (1998). Since the Sargan/Hansen-Test was rejected in most of the System GMM estimations, we only present the results for the Difference GMM estimator.

[19] The Sargan and Hansen Test on the joint validity of the instruments failed to pass the test for lags dated prior to t-6.

Table 3.5: GMM estimation of the average and the relative average skill Level of gross migration Flows

Dependent variable: average observable skill of a migrant (S_{kjt})				
	Average skills (1)	Average skills (2)	Relative average skills (3)	Relative average skills (4)
Δ Mean wage ($\Delta\mu_w$)	-0.186	-0.115	-0.195	-0.179
	(-1.52)	(-0.68)	(-1.60)	(-1.05)
Δ Mean empl. ($\Delta\mu_e$)	0.094**	0.066	0.090**	0.035
	(2.16)	(1.05)	(2.11)	(0.56)
Δ Wage dispersion ($\Delta\sigma_w$)	0.034	0.044	0.030	0.050
	(1.24)	(1.39)	(1.10)	(1.59)
Δ Empl. dispersion ($\Delta\sigma_e$)	0.031**	0.030**	0.022*	0.025*
	(2.45)	(2.05)	(1.72)	(1.71)
L.Mean Δ wage ($\Delta\mu_w$)		-0.012		0.061
		(-0.09)		(0.44)
L.Mean Δ empl. ($\Delta\mu_e$)		0.001		0.027
		(0.04)		(0.88)
L.Wage Δ dispersion ($\Delta\sigma_w$)		0.023		0.015
		(0.91)		(0.60)
L.Empl. Δ dispersion ($\Delta\sigma_e$)		0.021		0.017
		(1.43)		(1.19)
N	6318	5616	6318	5616
Sargan (p-value)	0.160	0.102	0.264	0.189
Hansen (p-value)	0.291	0.260	0.437	0.436
AR1 (p-value)	0.000	0.000	0.000	0.000
AR2 (p-value)	0.355	0.464	0.226	0.477
Instruments	39	38	39	38

Notes: robust t-statistics in parentheses; Significance levels: * 10%, ** 5%, *** 1%; GMM estimations are one-step estimates.

measure in Column (3). When adding lags of the independent variables, the significant effect of mean employment disappears, while the contemporaneous employment dispersion continues to attract better skilled migrants on average. Note, however, that all lags are insignificant, thus suggesting that lagged responses to regional income differentials are not relevant such that Models (1) and (3) remain the preferred specifications.

Overall, the robustness checks confirm our previous findings. Interestingly, it seems to be more important to take into account time-constant interregional differences that seem to strongly bias cross-sectional estimations than taking care of the potential endogeneity of the regressors due to reversed causality.

3.7 Conclusion

This paper examined the factors driving the skill selectivity of internal migration by proposing a framework for skill selective migration that takes account of the returns to skills in terms of both wages and employment. In a European context with strong and persistent employment disparities and with employment rather than wages responding to shocks, regional employment rather than wage disparities might be the driver of skill-selective migration. Our income-based model of skill selective migration predicts the average skill level of a migrant to be a positive function not only of regional differentials in wage inequality as suggested by the wage-based framework proposed by Borjas et al. (1992), but also of differentials in mean wages and mean employment rates as well as in employment inequality, i.e. differences in how employment chances are spread across the workforce.

Being able to exploit the variation in the skill composition of 702 gross labour flows across a ten-year period in Germany, we test these predictions based on a labour-flow fixed effects model that takes account of time-constant, flow-specific unobservables such as amenity differentials. This way, our model identifies the effects of interregional differences in the wage and employment distribution on migration selectivity beyond other relevant time-constant regional disparities. Comparing the outcomes to pooled OLS indicates that controlling for time-constant unobservables on the flow level is important in order to prevent biased estimates. This puts doubt on the reliability of previous cross-sectional estimates.

The findings suggest that regional employment differentials turn out to be important. In particular, a region attracts an increasingly skilled inflow of migrants, the higher its average employment. The same is true for an increasing employment inequality. The more unequal employment is spread across the regional workforce, the more a region attracts an increasingly skilled inflow of migrants. In contrast, regional differentials in the wage distribution exert no significant effect on the skill composition of labour flows. In a context where employment rather than wages tend to respond to regional shocks and resulting employment disparities are quite persistent, the allocation of human capital across space is thus driven by employment rather than wage disparities. Our paper thus establishes a missing link in understanding the self-reinforcing nature of interregional employment disparities. Using the flow between Eastern and Western

Germany as an example, the paper demonstrates that employment disparities have a much better predictive power for the observed skill composition between both parts of the country than wage disparities. However, wage disparities in response to regional shocks might gain in importance since institutions that give rise to regional wage rigidities in Germany such as collective wage agreements have started to erode (Ochel, 2005).

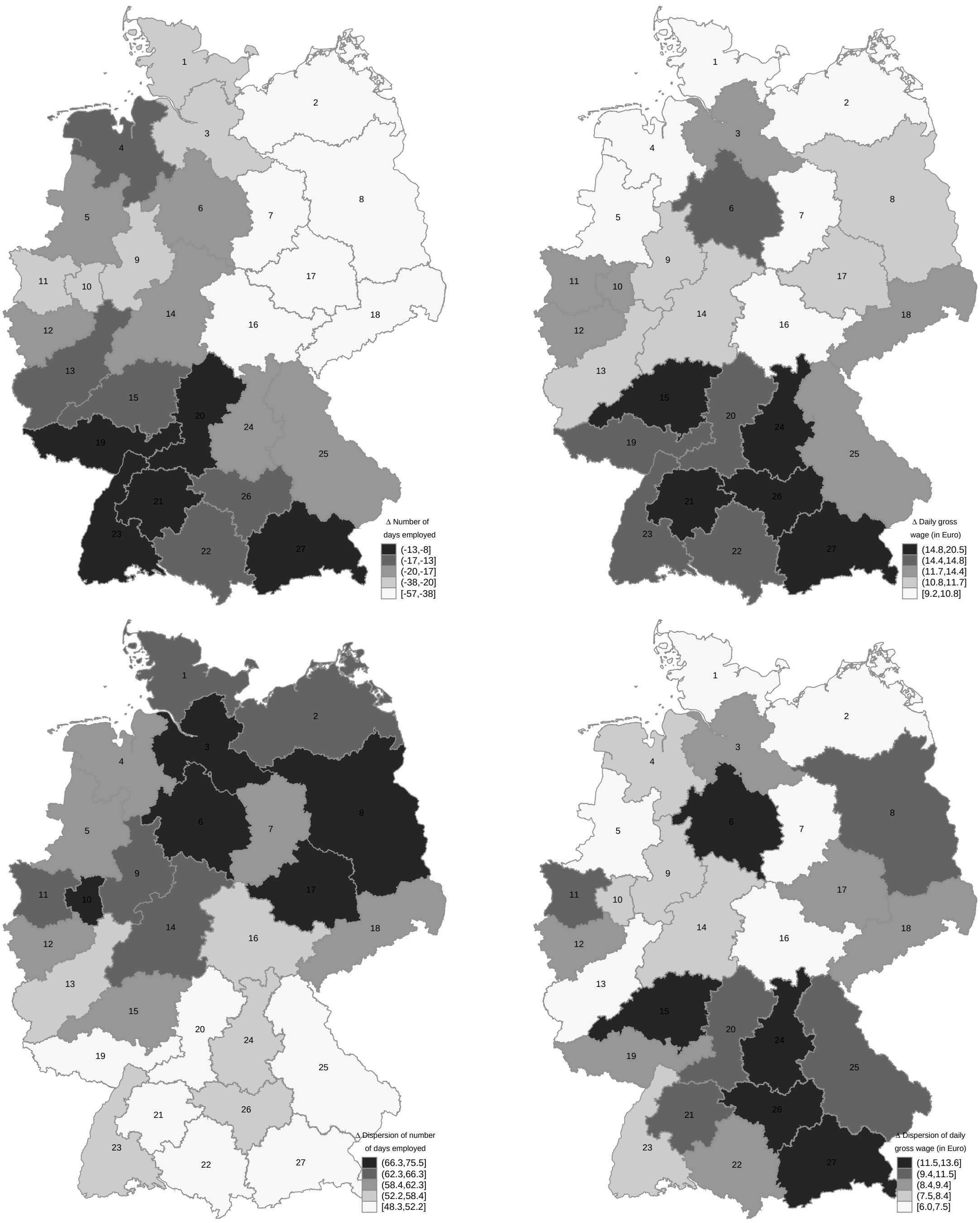
From a policy perspective, the results indicate that attempts to control migration flows in order to prevent an extensive brain drain should not focus on wage policies alone. In fact, attempts to artificially speed up wage convergence, as has been the case in Eastern Germany in the years following reunification, are likely to increase unemployment rates, thereby again fostering an increased brain drain. The average East-West migrant, for example, became more skilled relative to the average West-East migrant, thus aggravating the already existing brain drain with regard to observable skills.

The implications also apply to European regions that are particularly characterized by increasing employment disparities. As noted by Puga (2002), regional employment inequalities in Europe are increasingly driven by disparities within rather than between countries. During the recent recession, cross-country disparities have also been increasing dramatically. As a result, intra-European labour flows have already been diverted towards prospering high employment countries such as Germany (Bertoli et al. 2013). Given the insights of our paper, such flows are likely skill-biased with a brain drain from Southern to Northern European regions, thus potentially aggravating the current North-South divide. Therefore, Europe needs to aim at reducing regional employment disparities in order to mitigate related polarisation tendencies among its member states.

Finally, note that the predictions regarding skill selective migration might become more complex than suggested by our framework. In particular, routine and codifiable jobs typically ranking in the middle of the wage distribution have recently been found to experience lower wage and employment growth than the tails of the wage distribution (for a recent overview, see Autor 2013). Moreover, Autor and Dorn (2013) show that a corresponding wage and employment polarisation mainly takes place in local labour markets that initially specialized in routine tasks. As a consequence, interregional utility differentials may increase for both tails of the wage distribution at the same time suggesting that a region may be increasingly attractive for both low- and high-skilled workers, a prediction that is incompatible with a Borjas type framework of skill selectivity. However, this prediction only holds if we really observe polarisation patterns. For Germany, Senftleben and Wielandt (2012) show that wage and employment gains are mainly limited to the upper tail of the wage distribution which would predict increasing high-skilled migration compatible with our theoretical model. Hence, the relevance of recent polarisation tendencies for skill selective migration is far from resolved and a promising route for future research.

3.A Appendix

3.A.1 Absolute changes of the regional wage and employment distribution between 1995 and 2004



Part II

Minimum Wage Effects Along the Wage Distribution

4 The Minimum Wage Affects Them All: Evidence on Employment Spillovers in the Roofing Sector[1]

Joint with Bodo Aretz[2] and Melanie Arntz[3]

Abstract: This paper contributes to the sparse literature on employment spillovers on minimum wages. We exploit the minimum wage introduction and subsequent increases in the German roofing sector that gave rise to an internationally unprecedented hard bite of a minimum wage. We look at the chances of remaining employed in the roofing sector for workers with and without a binding minimum wage and use the plumbing sector that is not subject to a minimum wage as a suitable benchmark sector. By estimating the counterfactual wage that plumbers would receive in the roofing sector given their characteristics, we are able to identify employment effects along the entire wage distribution. The results indicate that the chances for roofers to remain employed in the sector in Eastern Germany deteriorated along the entire wage distribution. Such employment spillovers to workers without a binding minimum wage may result from scale effects and/or capital-labour substitution.

Keywords: minimum wage, Germany, capital-labour substitution, labour-labour substitution, scale effect

JEL-Classification: J38, J21, J23

[1]Published in *German Economic Review*, 2013, Vol. 14, Issue 3, pages 282-315.
[2]Centre for European Economic Research (ZEW), Mannheim.
[3]Universtity of Heidelberg and Centre for European Economic Research (ZEW), Mannheim.

4.1 Introduction

Most minimum wage (in short MW) research focusses on the average employment effect that MWs exert on workers with a binding MW, i.e. workers whose wage has to be raised in order to comply with the MW level. In a competitive labour market with a heterogenous workforce and an elastic product demand, for example, workers for whom the MW raises labour costs are expected to experience negative employment outcomes[2](Brown, 1999). However, depending on the production technology, the MW may also affect workers for whom the MW is not binding, see e.g. Neumark and Wascher (2008). If workers with and without a binding MW are complements, a negative scale effect that results from a reduced product demand negatively affects all workers' employment chances. If the two types of workers are substitutes, the MW may raise the demand for workers who earn a wage above the MW, thereby counteracting the negative scale effect by a positive substitution effect. In this case, we may observe negative employment effects for workers with a binding MW and even positive employment effects for workers with a non-binding MW. Moreover, profit-maximising firms may potentially substitute capital for the relatively more expensive labour input, thereby inducing an additional employment decline for all workers who are substitutable by capital. In this latter case, a firm might, in fact, lay off the poorest performers of each type of worker and reduce employment also among workers with a non-binding MW.

The existing literature mainly discusses employment spillovers, i.e. indirect employment effects for workers for whom the MW is not binding, as a potential source of bias. Linneman (1982), Currie and Fallick (1996), Abowd et al. (2000), and Neumark and Wascher (2000), e.g., identify the average employment effect on workers with a binding MW by comparing workers with and without a binding MW. Attempts to estimate substitution effects between workers tend to focus on the elasticity of substitution between skill or age groups rather than between workers with and without a binding MW.[3] The only study that we are aware of that focuses on employment effects along the wage distribution is by Neumark et al. (2004). They report evidence for a negative employment spillover for workers with a wage just above the MW level.

[2]In case of a monopsonistic labour market that allows employers to set wages below the equilibrium wage, a MW may instead induce positive or zero employment effects.

[3]See e.g. Neumark and Wascher (1995) for the substitution between age groups, Abowd and Killingsworth (1981), and Neumark and Wascher (1994) for the substitution between skill groups, and Hsing (2000) for substitution between part-time and full-time work.

The aim of this paper is to contribute to the sparse literature on employment spillovers by investigating employment effects along the entire wage distribution. In particular, our contribution is fourfold. First of all, we are able to analyse employment effects in a context where the MW bites very hard: the roofing sector in Germany. Its MW was introduced in 1997 and was subsequently raised several times. With a Kaitz Index, i.e. the ratio of the MW level and the median wage, that is around 1 in Eastern Germany, the bite has to be considered exceptional even by international standards (Machin et al., 2003; Dolton and Bondibene, 2011). The German roofing sector, thus, is an ideal setting to study employment effects along the entire wage distribution since its bite is likely to render indirect employment effects for workers above the MW. Our contribution thus extends the sparse literature on employment effects of MWs in the German construction sector for which König and Möller (2009) report negative evidence for Eastern Germany and positive but not always significant effects in Western Germany whereas Frings (2013) does not find significantly negative employment effects neither for western nor for Eastern Germany.

Secondly, we are able to exploit a natural experiment since, for institutional reasons, the MW was introduced only in parts of the construction sector including the roofing sector. Uncovered, yet comparable, sub-sectors may thus serve as a benchmark for the counterfactual development in the roofing sector in order to derive the average treatment effect on the treated (ATT) with respect to the chances of remaining employed in the roofing sector. Since the entire construction sector experienced a dramatic decline in demand after the end of the unification boom in the mid 1990s that almost halved the workforce in Eastern Germany, this is a highly relevant employment outcome.

Thirdly, we contrast the ATT from an intersectoral comparison with an ATT derived from a comparison of workers with and without a binding MW within the roofing sector. Under a number of identifying assumptions, a deviation between these ATTs may hint at employment spillovers within the roofing sector. In order to make such spillovers visible, we then combine both identification strategies. For this purpose, we estimate the counterfactual wage that workers of the control sector would receive in the roofing sector given their characteristics. This enables a comparison of workers with and without a binding MW across sectors and also allows for estimating the employment effects along the entire wage distribution.

Finally, we make use of two administrative linked-employer-employee panels one of which contains the full sample of workers in the roofing sector over the observation period of interest. This allows for approximating the hourly wage information in a much more precise way than in other studies dealing with the German construction sector such as König and Möller (2009) and Frings (2013).[4] Moreover, we are able to take account of unobserved heterogeneity at the individual level, which may be relevant if employers mainly substitute workers along unobservable skills as is suggested by Fairris and Bujanda (2008). Our paper, thus, yields much broader insights into the employment effects of MWs than most previous studies.

The findings indicate that the chances to remain employed in the roofing sector have deteriorated due to the MW introduction, especially in Eastern Germany where the bite of the MW was particularly hard. However, the impact suggested by comparing workers with and without a binding MW appears to be underestimated compared to the intersectoral comparison, thus hinting at employment spillovers of the MW on workers without a binding MW. An intersectoral comparison suggests negative employment outcomes for East German workers along the entire wage distribution. According to personal interviews with sector insiders, capital-labour substitution rather than scale effects drive this finding. Our results highlight the need for a broader perspective on the employment impact of MWs and also put doubts on any attempt to identify employment effects of MWs by comparing workers with and without a binding MW within a covered sector.

The paper is structured as follows. Section 4.2 contains information on the German roofing sector, the introduction of the MW and discusses some expectations for the empirical estimations given its market structure. Section 4.3 describes the data basis before Section 4.4 discusses the bite of the MW. The results on the average employment effects and the employment effects along the wage distribution are described in Section 4.5. Section 4.6 includes some robustness checks of the results before Section 4.7 concludes.

[4] The study by Rattenhuber (2014), who investigates the wage effects of the MW introduction in the construction sector, uses exact hourly wage information from the German salary and wage survey (Verdienststrukturerhebung). However, the data set only includes companies with more than 10 employees and does not comprise information beyond the year 2001.

4.2 The German Roofing Sector

Market structure. The goods and services that are provided by the roofing sector encompass the roofing of new buildings as well as the mending of old roofs. Roofing is a traditional craft in Germany requiring a master craftsman's diploma in order to start a business.[5] These traditional roofing companies usually employ less than ten employees and provide their services regionally and mainly to private home owners whose demand may be rather inelastic given the few available and mainly illegal substitutes such as moonlighting. In a survey among 250 roofing companies in 2011, more than three quarters of all companies considered quality rather than prices to be the main dimension of competition (Aretz et al., 2011). For those companies with more than 30 employees, which constitute less than 10% of all roofing companies, however, price competition may be more relevant since they tend to work for public contractors and are active beyond regional boundaries.[6]

Moreover, in contrast to most sectors that have been studied extensively in the MW literature, the roofing sector has a rather high level of qualification and is not very labour intensive. More than 95% of all workers work fulltime, and a relatively high share of around three quarters has at least a vocational training degree.[7] Moreover, labour costs account for less than 40% of total costs only (Cost Structure Survey 2001), and technical advances regarding materials and roofing techniques appear to be quite important as reported by roofing companies in a number of qualitative interviews.[8]

Business cycle. The entire construction sector experienced a boom period in the early 1990s due to German reunification but began to shrink from the mid 1990s on, see Figure 4.1. In Eastern Germany, this post-unification downturn was much more dramatic than in Western Germany and reduced the construction sector's revenues in the subsequent years by more than half. After 2004, all construction sectors reinstalled revenue levels in Western Germany similar

[5] As an exception, it is not required to hold such a diploma if someone works as an itinerant worker. Such workers tend to work alone and mainly provide mending services only.

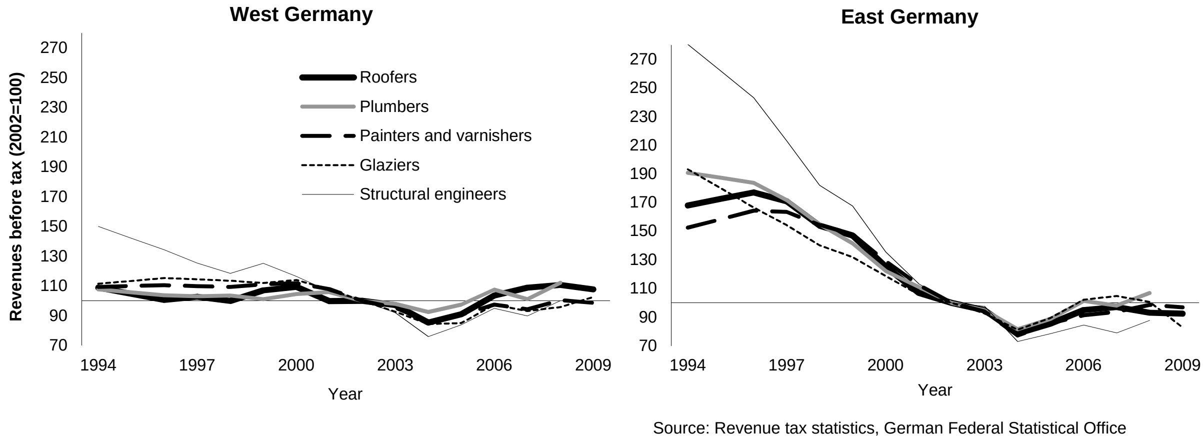
[6] Information on company size is based on the BA data (see Section 4.3 for details).

[7] The part-time information is taken from the Cost Structure Survey for 2001 (Kostenstrukturerhebung), which is released by the German Statistical Office. The share of qualified workers is calculated based on the BA data (see Section 4.3 for details).

[8] Ten qualitative interviews with roofing companies and four additional interviews with representatives of the trade union and the employer's association were conducted within a report prepared for the Federal Ministry of Labour and Social Affairs, see Aretz et al. (2011) for details.

to the early 1990s, while the recovery in Eastern Germany was rather marginal. Compared with structural engineering, the roofing sector and other sub-construction sectors such as plumbing, glazing and painting services experienced a less dramatic decline in the demand for their services in the mid 1990s and a faster recovery after 2004. The demand for sub-construction work hinges on the demand for new buildings as well as the age structure of the existing stock of houses with the latter apparently having a smoothing impact on the business cycle compared to structural engineering.

Figure 4.1: Overall revenues in Western and Eastern Germany by sector, 1994 - 2009



Moreover, sub-construction sectors broadened their portfolio during the last years, thereby stabilising the demand for their services. In particular, roofing companies are increasingly involved in the assembling of photovoltaic cells as well as the ex post insulation of old roofs.[9] The plumbers and, to a lesser extent, glaziers and painters also benefited from this development. At least in Western Germany, this has presumably contributed to a faster recovery in the roofing and the plumbing sector compared to the other sub-construction sectors and structural engineering.

Minimum wage regulations. Apart from shrinking demand, additional pressures in the mid 1990s stemmed from the introduction of a free movement of labour that allowed Eastern European firms to send workers to German construction sites while paying home country wages. In order to protect German workers, legally binding MWs that had to be paid to all workers on German construction sites irrespective of the origin of their contract were introduced in the structural engineering and some sub-construction sectors. Since MWs are negotiated

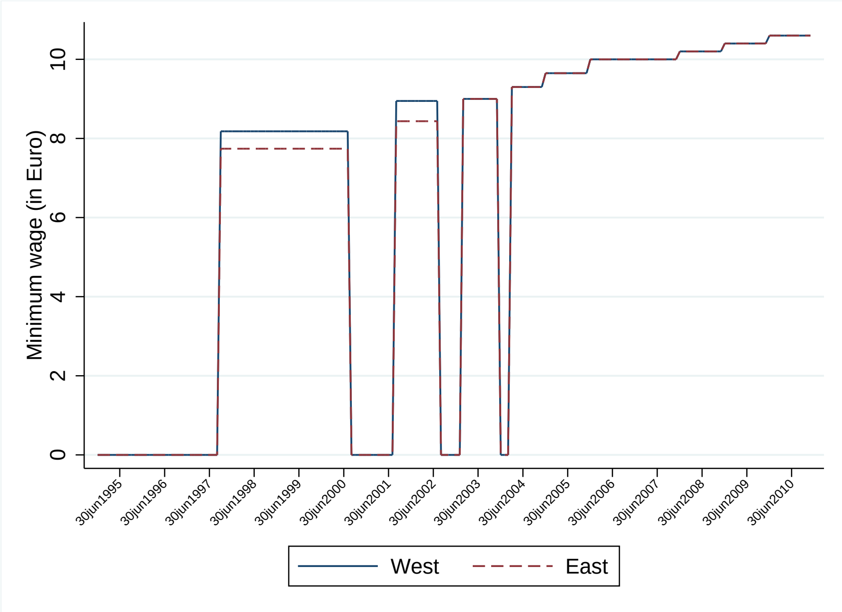
[9]Both of these developments have been boosted by government initiatives for subsidising solar energy generation since 2000 (*Erneuerbare-Energien-Gesetz*) and energy-saving renovations since 2002 (*Energetische Gebäudesanierung*).

separately for certain sub-divisions of the construction sector, not all sub-divisions agreed on MW regulations, resulting in a coexistence of quite comparable sectors with a legally binding MW (e.g. structural engineering and roofing sector since 1997; painting sector since 2003) and sectors such as glazing and plumbing services that are not subject to a legally binding MW until now. Hence, these sectors may potentially serve as a benchmark for the counterfactual development in the roofing sector in the absence of a legally binding MW.

The MW in the roofing sector applies to all blue-collar workers of any roofing company or roofing branch within a larger company who are at least 18 years of age, who are not an apprentice and who are not working as a custodial worker. Thus, all white-collar workers such as office clerks as well as certain parts of the blue-collar workforce are exempted from the MW regulation. Introduced in October 1997, the MW was subsequently raised several times but was also interrupted by short periods without any legally binding MWs, see Figure 4.2. These interruptions reflect the fact that the MW is negotiated between the responsible trade union (*IG Bau*) and the association of employers in the roofing sector (*Zentralverband des Deutschen Dachdeckerhandwerks*) as a part of the general collective bargaining agreement. When these agreements expire, there may be short interruptions before a new agreement is reached. Because the continuation of a MW was not subject to any debate since its introduction, roofing companies could, however, expect a new MW agreement, rendering any behavioural adjustments during these interruptions very unlikely. Moreover, MWs were harmonised between western and Eastern Germany in 2003 despite wages in Western Germany exceeding wages in Eastern Germany by about 25%. This results in an extremely hard bite of the MW in Eastern Germany as we will see in Section 4.4.

Taking all this evidence together, the roofing sector's market structure suggests a rather limited impact of MWs on employment given its limited labour intensity, the ability of roofing companies to at least absorb some of additional costs by raising prices and the fact that technical advances and increases in productivity offer options for cushioning rising labour costs. At the same time, however, the lower wage floor was fixed on a rather high level (see also section 4.4), particularly in Eastern Germany, thus rendering employment effects likely. Moreover, changes in relative input prices may create incentives for substituting labour by capital and/or less skilled by skilled workers. Finally, the MW in the roofing sector was introduced during a period of economic downturn and a shrinking market size. This strongly reduced the sector's workforce

Figure 4.2: Minimum wage level in the German roofing sector by region, 1995-2010



(see Appendix 4.A.1), although the number of companies even slightly increased at the same time as the share of single-person companies jumped from 8% in 1995 to 23% in 2010. With the number of unemployed workers with sector-specific human capital queuing for jobs on a rise, the bargaining power of those still working in the sector may have come under pressure.

4.3 Administrative Linked Employer-Employee Data

For our analysis, we are able to exploit two administrative linked employer-employee panel data sets: (1) data that is collected by the central pay office of the roofing sector (*Lohnausgleichskasse, LAK*), in short the LAK data, and (2) data that is collected by the Federal Labour Agency (*Bundesagentur für Arbeit, BA*) for all employees that are subject to social insurance contributions, in short the BA data.

4.3.1 LAK Data

In order to balance out the seasonal fluctuation of the sector, all roofing companies have to pay an insurance premium to the LAK that is related to the total payroll of their blue-collar workers. Therefore, they are obliged to give a monthly record to the central pay office of the roofing sector. For our analysis, we have access to the full sample of blue-collar workers on a monthly basis for the years 1995 to 2010, thus covering both the pre- and post-MW period. Information on monthly working hours and monthly gross wage allows for calculating the hourly gross wage. Between October and April, however, reported working hours need not match the true working hours because of special regulations for cushioning the seasonal character of the sector's activities. Hence, we use the June information for the analysis based on the LAK data in order to avoid such distortions and to ensure the comparability of the analysis with the BA data (see below).

The LAK data contains additional information only on sex and age of the workers. Since we do not know whether someone is an apprentice or working as a custodial worker, who are both exempted from the MW regulation, we are not able to exactly identify all covered workers. Since most custodial workers are female, however, and the share of females among covered roofers is less than 2% according to the BA data, we exclude women from the LAK sample. We also exclude all workers below the age of 19 and assume that this also eliminates most uncovered apprentices. Our sample, thus, differs from the exact coverage by missing some covered women and including some uncovered apprentices in the sample. Overall, we observe a total of 1,094,609 observations between 1995 and 2010 that stem from 217,779 individuals in 22,879 firms. Note that we are able to calculate some firm level information such as average gross pay, average firm size and average age of the company's workforce that we can use in addition to the individual information.

4.3.2 BA Data

A major disadvantage of the LAK data is that it is only available for the roofing sector, thus precluding any identification strategy that rests upon intersectoral comparisons. Such an alternative identification strategy, however, becomes available based on the BA data since it includes information for 75% of all companies in the roofing sector as well as sub-samples of companies in other sub-construction sectors such as painting, plumbing and glazing services for the observation period from 1994 to 2008.[10] For all individuals who are subject to social insurance contributions and work in one of these companies on June 30th, the data contains the corresponding period of continued employment in that company within the calender year

[10]This information is taken from the *Betriebshistorikdatei*, a data set that aggregates the individual data that is collected by the BA to the firm level, see Hethey-Maier and Seth (2010) for details on the data.

that overlaps June 30th.[11] Thus, the longest spell encompasses the full calender year, while the shortest employment spell would be an employment period of one day on June 30th only.

For each employment spell, we have information on age, sex, educational background, the gross daily wage, occupation, and occupational status. Thus, the data allows for identifying covered individuals quite precisely. In particular, we are able to exclude custodial workers, apprentices and white-collar workers as well as underage workers.[12] Overall, the sample consists of 791,910 observations in the roofing sector that stem from 172,257 covered roofers in 17,186 roofing firms and 1,557,661 observations by 354,834 workers in 35,250 firms from other sub-construction sectors who fulfill the same criteria.

Since the data only distinguishes between full-time and part-time workers and includes information on daily gross wages only, the main restriction of the BA data refers to the corresponding lack of information on hourly gross wages. As a remedy, we impute the hourly gross wage by estimating the observed hourly gross wage in the LAK data as a function of explanatory variables that are available in both data sets. For this purpose, we first adjust the LAK data to have a similar data structure as the BA data by creating employment spells for each individual who has worked on June 30th. For these spells, both data sets provide information on or allow for computing the length of the spell, the beginning of the spell, the daily gross wage, dummies for part-time or full-time employment, individual information on sex and age as well as a number of firm-level information such as firm size, workforce composition, and average gross daily wage.[13] Using all these explanatory variables and allowing for additional heterogeneity by estimating the wage model separately for each year, Eastern and Western Germany as well as for workers of different quintiles of the daily gross wage distribution, we are able to explain 88% of the variation in hourly gross wages in the LAK data. We then use these estimates for predicting the hourly wage in the BA data. The quality of this imputation not only hinges on the R^2 of the wage estimation, but also depends on the comparability of the LAK and the BA sample and explanatory variables. Appendix 4.A.2 shows that the imputed and observed wage distribution are very comparable. As a result, the average predicted mean wage for full-time workers of

[11] This information is taken from the employee record of the BA (*Beschäftigtenmeldungen*), see e.g. Drews (2008) for details.

[12] We also exclude workers with a minor employment which is defined as earning below the social insurance contribution threshold of 400€ per month because these workers are included in the data only after 1998. For comparability reasons, we also dropped such workers in the LAK data.

[13] Although we do not use women in the LAK analyses, we do use their LAK wage and impute the corresponding wage in the BA data to be able to include them in the analyses conducted with the BA data.

13.26€ and 9.94€ in the BA data in western and Eastern Germany, respectively, comes very close to the observed average wage in the LAK data with 13.22€ for Western Germany and 9.85€ for Eastern Germany.[14]

4.4 The Minimum Wage and Its Bite

Table 4.1 displays several indicators of the bite of the MW for the June preceding the introduction of a new MW regulation within the next year. In particular, we look at the share of covered workers for whom the upcoming MW is binding due to earning a wage below the minimum in the June preceding the new MW regulation.[15] We also show the average wage increase these workers would have to receive in case of full compliance with the upcoming MW. This individual wage gap for a worker i with a binding MW in period t is thereby defined as follows:

$$wage\ gap_{it} = \frac{MW_{i,t+1} - w_{it}}{w_{it}}, \tag{4.1}$$

where w_{it} represents the workers hourly wage and $MW_{i,t+1}$ the upcoming MW. We contrast this wage gap to their actual wage increase within the next year and the actual wage increase during the same time period among workers for whom the MW was not binding. We complement this information by the Kaitz-Index, i.e. the ratio between the MW level and the median wage in the sector. Note that the indicators may slightly underestimate the bite of the MW due to the fact that the hourly wage may contain overtime compensation that is not subject to the MW.[16]

The indicators based on the LAK data allow for several interesting insights. First of all, the share of covered workers for whom the MW was binding by the time of its introduction was as low as 1.3% in Western Germany compared to 12.5% in Eastern Germany. While this share rose up to 5% in Western Germany until 2008, around 50% of all eastern workers earned below the upcoming MW at that time, a share that clearly exceeds the impact level that Machin et al. (2003) considered a hard biting MW. This extreme bite was fostered by the introduction of a

[14] Thus, the hourly wage information in the BA data is approximated in a much more precise way than in other studies, e.g., by König and Möller (2009) and Frings (2013).

[15] We do not adjust for nominal wage changes between the two dates of comparison because the intermediate time span is quite short.

[16] On average, overtime hours account for 6% of the working hours in June and, thus, may lead to an estimated hourly wage that is up to 1.6% too high depending on the applied overtime compensation scheme ranging from no additional compensation to a markup of 25%. Since we do not know which scheme is applied, we left the data uncorrected as the resulting imprecision appears to be rather marginal.

Table 4.1: Indicators of the MW bite measured in June prior to the next MW regulation, LAK and BA data

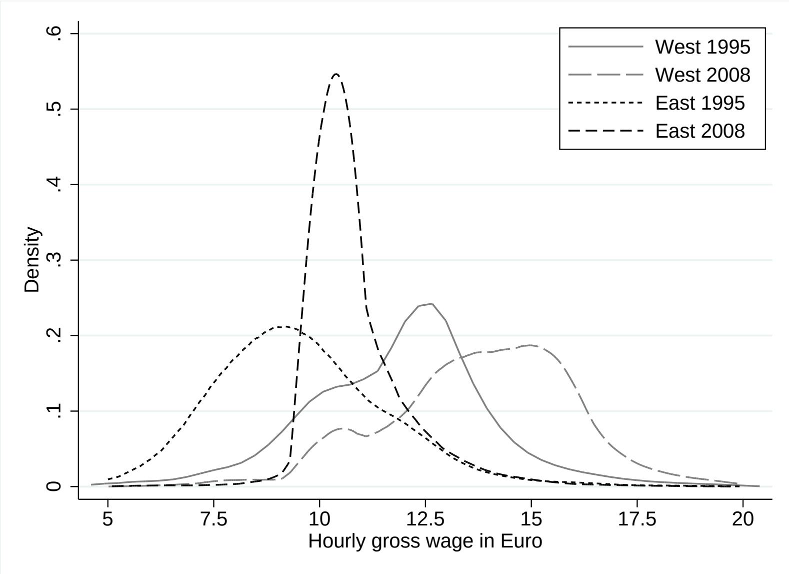
New MW regulation takes effect on	MW (in Euro)	Workers with a binding MW? Yes: Share (in %) LAK	Yes: Share (in %) BA	Yes: Wage gap[a] (in %) LAK	Yes: Δ Wage[b] (in %) LAK	No: Δ Wage[b] (in %) LAK	Kaitz Index LAK
Western Germany							
01.10.97	8.2	1.3	2.4	11.0	7.2	2.3	64.7
01.09.01	9.0	1.5	3.9	8.7	6.8	1.4	67.2
01.03.03	9.0	1.5	3.4	8.9	6.0	2.4	67.2
01.04.04	9.3	2.2	4.8	8.1	5.7	1.4	68.4
01.05.05	9.7	2.9	5.8	8.5	4.9	0.6	70.3
01.01.06	10.0	4.4	6.9	7.9	4.9	1.1	72.6
01.01.07	10.0	4.6	7.5	8.1	6.7	3.2	72.7
01.01.08	10.2	5.4	8.2	6.7	5.3	2.2	73.1
01.01.09	10.4	4.9	7.5	6.6	8.1	3.0	73.4
Eastern Germany							
01.10.97	7.7	12.5	11.5	9.7	6.7	0.0	82.0
01.09.01	8.4	14.2	12.0	3.9	4.6	0.6	89.2
01.03.03	9.0	34.1	23.3	4.2	4.1	0.1	95.0
01.04.04	9.3	44.1	28.7	3.8	4.1	0.3	97.9
01.05.05	9.7	46.9	33.5	4.3	4.0	0.1	99.2
01.01.06	10.0	55.5	40.8	4.1	4.0	0.1	100.2
01.01.07	10.0	45.5	28.1	1.6	1.9	0.9	99.6
01.01.08	10.2	53.5	32.1	2.6	3.3	1.3	100.7
01.01.09	10.4	50.0	28.9	2.4	3.3	0.7	99.9

[a] Wage gap refers to equation (4.1)
[b] Δ wage corresponds to the actual observed percentage wage change $(w_{it+1} - w_{it})/w_{it}$ between the June preceding and the June following the new MW regulation.

common MW level in both parts of the country in 2003. Since then, the MW level approximately corresponds to the median wage in Eastern Germany so that the Kaitz-Index ranges around 100%. Even in Western Germany, the Kaitz-Index still ranges between two thirds and three quarters of the median wage. Compared to Dolton and Bondibene (2011), who find the Kaitz-Index to range between 30% and 70% in a survey among 22 OECD countries, the bite of the MW in the roofing sector is thus extremely hard, especially in Eastern Germany.

We also observe that the MW has been effective, i.e. actual wage increases among workers with a binding MW exceeded the wage increases among workers for whom the MW was not binding. While the change in the west German wage distribution is rather marginal, the wage compression in Eastern Germany results in a huge spike of workers whose wages range around the MW level, see Figure 4.3. Finally, note that despite these actual increases, they still fall short of the increases workers would have had to receive in case of full compliance, especially in

Figure 4.3: Kernel densities of hourly gross wages in Eastern and Western Germany, 1995 and 2008, LAK data



Western Germany during the initial years after the MW introduction. The improved compliance with the MW regulation during the last years might be due to stronger controls after 2006 according to interviews that we conducted with sector insiders.

Table 4.1 also contrasts the share of workers with a binding MW based in the LAK data to the corresponding share based on the imputation in the BA data. In contrast to the LAK data, however, the share of workers with a binding MW follows a probabilistic concept because we do not only impute the mean wage prediction for each individual but also the corresponding distribution that results from the unexplained variance and the variance of the estimated parameters. Assuming this distribution to be normally distributed, we are then able to calculate the probability that the wage of a worker falls below the MW level which we denote by P_{MW}. For the BA data, Table 4.1 thus reports the average predicted probability of being affected by a binding MW among all covered workers.

As we can see, the resulting share of workers with a binding MW resembles the LAK patterns but differs in levels, especially for Eastern Germany in the last years. In fact, imputing the probability of being affected by a binding MW in the LAK data, gave very similar deviations to the observed share of workers with a binding MW. Hence, it is apparently the extreme wage compression that leads to the asymmetric form of the wage distribution in Figure 4.3 and, thus, to a systematic underestimation of the share of workers with a binding MW in Eastern Germany, see Appendix 4.A.3 for further explanation.

Despite the large bite of the MW, especially demonstrated by the high wage compression in Eastern Germany, the overall labour cost burden is still modest for two reasons. First, even in the case of full compliance with respect to the MW regulations, total labour costs only increased by 1% in Western Germany and 2.5% in Eastern Germany on average during the observed time period. Second, labour costs amount to less than 40% of total costs, so that the change in average total costs varies across time between 0.2-0.5% in Western Germany and 0.3-0.8% in Eastern Germany. However, despite the low impact on total costs on average, some firms may well be affected more strongly. Moreover, the cost burden may cumulate over time due to the gradual increase in the MW level.

Finally, it is worth mentioning that individuals with a binding MW clearly differ between western and Eastern Germany. While the average worker with a binding MW in Eastern Germany does not differ much from an average worker without a binding MW, the average worker with a binding MW in Western Germany rather corresponds to a marginal worker with below average human capital, short tenure and part-time employment in firms with a skill and wage level below average, see Table 4.2.

Table 4.2: Characteristics of workers in Western and Eastern Germany by binding status, BA data 1995-2008

	Western Germany		**Eastern Germany**	
MW for workers is binding?	Yes	No	Yes	No
Individual characteristics				
Worker with voc. training deg. (in %)	24.1	67.2	70.2	80.4
Workers without voc. training (in %)	34.5	31.1	25.2	19.0
Part-time workers (in %)	41.3	1.7	4.6	0.5
Previous work exp. in sector (in years)	2.2	4.3	2.9	3.6
Previous tenure in firm (in years)	1.9	3.7	2.2	3.0
Firm characteristics				
Average firm size	4.0	6.2	5.7	7.8
Firm's share of skilled workers (in %)	63.1	83.1	79.9	82.2
Firm's mean daily gross wage (in €)	51.78	72.33	50.43	56.08
Number of observations	15,523	485,640	39,960	196,981

Note: Workers with $P_{MW} > 0.5$ are considered to be bound by the MW.

4.5 Employment Effects

Since the MW was introduced for the entire sector at the same time, a strategy for the identification of the MW impact on employment cannot rest on regional variation as has been done in many US studies (Dube et al., 2007, 2010; Card and Krueger, 1994, 2000). Exploiting the existing variation in the MW level between Eastern and Western Germany in the mid 1990s is also not advisable since the business cycle after the reunification boom differed between both parts of the country, see Figure 4.1. Moreover, recent attempts to evaluate MWs using the incremental differences-in-differences framework (Dolton et al., 2011) do not seem able to convincingly isolate the causal effect of each MW increase from lagged or anticipation effects. Thus, there are mainly two potential approaches available for the identification of employment effects. Either one uses a sub-construction sector that is not covered by a MW regulation but is as similar as possible to the roofing sector (intersectoral comparison), or one exploits the variation in treatment intensity within the roofing sector by comparing workers with and without a binding MW (intrasectoral comparison).

We will apply both approaches using a difference-in-differences (DiD) framework and concentrate on measuring the chances of remaining employed in the same sector.[17] For this, let e_{it} denote the employment status in period t for an individual i. In particular, let $e_{it+1} = 1$ in case of being employed in the same sector as in the previous period and $e_{it+1} = 0$ otherwise, an outcome measure that we are able to observe in both the LAK and the BA data. As the outcome variable is binary, a logit estimation might be appropriate. However, we believe that it is important to consider individual fixed effects because selection on unobservable characteristics may be relevant. Since fixed effects can not be identified in a logit model framework (Wooldridge, 2002), we estimate a simple linear probability model (LPM) with fixed effects in the following way:

$$P(e_{it+1} = 1) = \alpha_g + \delta D_{it} + \beta X_{it} + \gamma_p + \nu_t + c_i + \epsilon_{it}, \tag{4.2}$$

[17]This outcome measure is of main interest in a market context that is dominated by a shrinking market size and a corresponding reduction in employment since the mid 1990s, compare Appendix 4.A.1. The question of whether someone was able to keep his job in this market context given the additional cost pressures of the MW is of main concern. On average, 80% (77%) of all western (eastern) German roofers are still employed in the same sector after one year. Note, that this outcome should not be equated with effects on the total employment in the roofing sector. As an example, additional market entries by single person companies are not captured by this employment outcome.

where α_g captures the time constant difference between the control and the treatment group, γ_p allows for changes in the period after compared to the period before the MW introduction that are common to both groups, ν_t captures year-specific effects, c_i depicts the individual fixed effects, and X corresponds to a set of control variables.[18] D_{it} is the treatment indicator with $D_{it} = 1$ for individuals of the treatment group after the treatment has taken place, and $D_{it} = 0$ otherwise.[19]

In Section 4.5.1 we show the results for the intersectoral and intrasectoral approach. If all underlying assumptions were valid, we should yield similar results for both approaches. As we will see and further discuss in Section 6, this is not the case, indicating that spillover effects of the MW on workers earning above the MW level are present. In Section 5.2, we therefore explicitly estimate the employment effects along the wage distribution by running the intersectoral comparison for workers falling in different wage deciles.

4.5.1 Average Employment Effects

Intersectoral Comparison

A feasible control sector needs to capture the counterfactual change in employment outcomes for roofers in the absence of the MW. For this to be a plausible assumption, the control sector should have a comparable market structure as well as comparable demand conditions. Among the sub-construction sectors without a legally binding MW - the plumbing and the glazing sector[20] - the plumbing sector is preferable for a number of reasons. According to Figure 4.1, the business cycle in the plumbing rather than the glazing sector resembles the business cycle in the roofing sector. In fact, for Western Germany demand conditions almost follow the same path, while in Eastern Germany the demand for plumbing services started to drop somewhat earlier than in the roofing sector, a deviation that we will return to in the robustness analysis in section 4.6.

[18] All estimations for the BA data include individual and firm covariates. On the individual level the occupational status and educational attainment (6 dummies) are used as explanatory variables. On the firm level we use age and qualification of the company workforce, company size (4 dummies), and 2nd order polynomial of the mean daily gross wage. The same covariates except educational attainment are used for the LAK data.

[19] We find that in most cases, only few observations have predictions outside the plausible range. We also estimated the models using pooled Logit and pooled OLS. The pooled results turned out to be quite similar indicating that the LPM model performs well; for more details see Aretz et al. (2012a).

[20] The painting sector introduced a MW in 2003.

Table 4.3: Comparison of the roofing, the glazing, and the plumbing sector by various economic indicators

	Roofers	Plumbers	Glaziers	Source
Number of companies	11,295	37,720	3,752	A, 1996
Number of employees	113,996	364,393	25,393	A, 1996
Avg. number of employees per company	8.8	9.0	6.6	A, 1996
Share of firms by revenues (in 1,000)				B, 1996
<100 DM	6.8	8.8	13.6	
100-500 DM	24.6	33.7	42.6	
500-$1,000$ DM	26.1	23.5	21.5	
$1,000$-$2,000$ DM	25.1	19.3	13.5	
$>2,000$ DM	17.4	14.6	8.5	
Value added in € per employee	37,195	35,949	32,931	C, 2001
Investments/employee (in €)	1,472	1,229	2,482	C, 2001
Share of labour costs (in %)	36.0	32.5	49.0	C, 2001
Avg. gross daily wage/fulltime employee (in €)	61.25	63.23	64.28	A, 1996
Number of companies/1 Mio. sector revenue	1.3	1.3	0.6	B, 1996

Note: A - BA data (see Section 4.3); B - Revenue tax statistics of the German Federal Statistical Office (Umsatzsteuerstatistik); C - Cost Structure Survey of the German Federal Statistical Office (Kostenstrukturerhebung)

According to Table 4.3, the plumbing sector is similar to the roofing sector with regard to important market indicators that moderate the potential impact of a MW.[21] In particular, roofing and plumbing companies are similarly sized in terms of both the number of employees and the revenues generated. Also, the value added is highest in the roofing sector, closely followed by the plumbing sector. Moreover, the glazing sector is more labour-intensive and invests almost twice as much per employee than the other sectors while the average gross daily wage is quite comparable across all sectors. Finally, the number of companies per one million euro of revenues in the sector, a measure of the degree of competition, is almost identical in the roofing and plumbing sector but much lower in the glazing sector, suggesting less competition.[22]

[21] We display the pre-MW indicators for 1996 wherever it is available as the basis for judging the usefulness of a sector as a benchmark for the roofing sector.

[22] Although Table 4.3 suggests that the roofing and the plumbing sector are very similar, a high comparability is not ensured by all means. The skill level, e.g., is higher in the plumbing sector as 18% of the roofers (plumbers) have no vocational education but only 6% of the plumbers.

Table 4.4: Average employment effects, inter- and intrasectional comparison, Eastern and Western Germany

	(1)	(2)	(3)	(4)
Comparison	**Intersectoral**		**Intrasectoral**	
Model specification	Pooled LPM	FE LPM	FE LPM	FE LPM
Data set	BA	BA	LAK	BA
Observation period	1994-2007	1994-2007	1995-2007	1995-2007
Eastern Germany				
ATT in pp[d]	-2.3***	-2.9***	-8.5***	-10.0***
Robust s.e.[c]	(0.3)	(0.3)	(0.7)	(0.8)
Obs. (in 1000)	497	497	288	224
Share of $\hat{Y} \notin [0;1]$	0.4%	9.0%	7.3%	8.6%
ATE for all covered workers	-2.3	-2.9	-1.7	-2.0
Western Germany				
ATT in pp[d]	1.2***	-1.2***	-2.7	-5.0*
Robust s.e.[c]	(0.2)	(0.2)	(2.3)	(2.0)
Obs. (in 1000)	1,110	1,110	601	457
Share of $\hat{Y} \notin [0;1]$	0.9	0.4%	0.1%	0.9%
ATE for all covered workers	1.2	-1.2	-0.1	-0.3

[a] DiD estimation for workers in the roofing and plumbing sector
[b] DiD estimation for workers with and without a binding MW within the roofing sector
[c] Standard errors are clustered at the individual level.
[d] All estimations for the BA data include as explanatory variables on the individual level the occupational status and educational attainment (6 dummies) in the fixed effects estimations plus age, age^2, sex, 2nd order polynomial of tenure and previous work experience in the sector in the pooled estimations. On the firm level, age and qualification of the company workforce, company size (4 dummies), and 2nd order polynomial of the mean daily gross wage are used. The same covariates except educational attainment are used for the LAK data.
Significance levels: * 5%, ** 1%, *** 0.1%

Therefore, we consider the plumbing sector as a suitable and better benchmark for the roofing sector than the glazing sector. For the intersectoral comparison, this means that the treatment group corresponds to all workers of the roofing sector that are covered by the MW regulations, while workers in the plumbing sector, who would have been covered if they worked in the roofing sector, are considered as the control group. In equation (4.2) this means that D_{it} equals one for all roofers in the period after the MW introduction (1997-2007) and zero otherwise. Hence, this approach can only be estimated based on the BA data. The treatment refers to being covered by the MW regulations and the resulting estimates give us the ATT in the roofing sector if changes in employment outcomes of plumbers between the ex-ante situation (1994-1996) and the ex-post situation (1997-2007) capture the counterfactual change in employment outcomes for roofers in the absence of the MW. Moreover, there may not be any control group contamination, i.e. there is no indirect effect of the MW regulations in the roofing sector on the plumbing sector.

Column (1) and (2) in Table 4.4 show the ATT for the average treated worker in Eastern and

Western Germany from a pooled LPM and from a LPM that takes account of individual-specific fixed effects (FE LPM) using standard errors clustered at the individual level.[23] Irrespective of the specification, the MW in Eastern Germany appears to have reduced the chances for roofers to remain employed in the sector by around 2 to 3 percentage points on average compared to plumbers who have not been subject to MW regulations. In Western Germany, the findings are not as clear. While the pooled LPM suggests a positive MW effect of 1.2 percentage points, the FE LPM indicates that the chances for a roofer to remain employed in the next year after the MW was introduced decreased by 1.2 percentage point compared to plumbers who have not been subject to MW regulations. This suggests that MWs in Western Germany increased layoffs mainly among workers with poor unobservable characteristics so that pooled estimations are upward biased.

Intrasectoral Comparison

Next, we conduct a comparison within the roofing sector. The treatment group in this approach corresponds to roofers with a binding MW due to earning a wage in June that is below the MW level that takes effect until June of the next year. For the pre-regulation years, we consider someone to belong to the treatment group if his wage falls below the MW level that would have to be applied in the pre-regulation years given the increases of the median nominal wage in the LAK data. Workers of the roofing sector whose wages are above that minimum level are used as the control group. While we are able to exactly identify these groups in the LAK data, we define the treatment group in the BA data to encompass all covered workers whose probability to fall below the MW level exceeds 50%.[24]

Hence, the treatment indicator D_{it} in equation (4.2) equals one for all covered workers with a binding MW after the MW introduction and zero otherwise. The estimates give the ATT if changes in employment outcomes for workers without a binding MW between the ex-ante situation (1995-1996)[25] and the ex-post situation (1997-2007) capture the counterfactual change

[23] Additional firm-level clustering in the pooled models using two-way clustering as proposed by Cameron et al. (2011) did not change the inference much. Within a fixed effects context, firm-level clustering is not feasible as individuals are not nested within firms because of inter-firm movements. Descriptive statistics regarding both the dependent variable and the set of covariates for both roofers and plumbers prior to and after the MW introduction are provided in Aretz et al. (2012a).

[24] A specification with P_{MW} as a continuous treatment variable gave similar results.

[25] Note that we have only two pre-regulation years because the LAK data that are used for imputing hourly wages are only available from 1995 onwards.

in employment outcomes for the treated roofers in the absence of the MW.

This approach can be estimated using both the BA and LAK data.[26] Columns (3) and (4) of Table 4.4 show the DiD estimates for the intrasectoral comparison. The results indicate that workers with a binding MW were 9 to 10 percentage points less likely to remain employed after the MW introduction relative to workers without a binding MW. This negative effect is confirmed by both the LAK and BA data suggesting that the BA data yield quite reliable estimates despite the imprecision in the distinction between workers with and without a binding MW. For Western Germany, the estimates based on the LAK data show an insignificant reduction in the probability of continued employment by -2.7 percentage points, while the treatment effect is slightly larger and significant at the 5% significance level when using the BA data.[27]

If all identifying assumptions regarding common trends in the absence of the treatment and the lack of any spillovers held, multiplying the ATT from the intrasectoral comparison by the share of workers with a binding MW should yield the ATT from the previous intersectoral comparison, i.e. the average treatment effect (ATE) for all covered workers in the roofing sector.[28] The implied ATEs are included in Table 4.4. For Eastern Germany, this yields an ATE of $-2.0(= -10.0 \times 0.204)$ percentage points for the BA results since, on average, 20.4% of all workers in Eastern Germany are affected by a binding MW across the entire period. In Western Germany, the MW is binding for 5.5% of the workforce on average, implying an ATE of -0.3% percentage points. Compared to the fixed-effects estimates from the intersectoral comparison in column (2) of Table 4.4, the fixed-effects estimates in column (3) and (4) appear to be underestimated. This deviation can be due to a violation of the common trend assumption and/or spillover effects between the control and treatment group in either the intersectoral and/or the intrasectoral approach. In fact, robustness checks in Section 4.6 indicate that spillover effects between workers with and without a binding MW are the main driving forces of these deviations as all other underlying assumptions seem to hold quite well. In a next step, we therefore explicitly estimate the employment effects along the wage distribution by running the intersectoral comparison

[26] Corresponding descriptives for the BA data on the intrasectoral comparison are provided in Appendix 4.A.4. Descriptives for the LAK data are available from the authors upon request.

[27] When including individuals with a minor employment in the LAK estimates, the treatment effect amounts to highly significant -17.3 percentage points. For a better comparability with the BA data, we leave these individuals out. Still, the estimates with minor employment indicate that the MW may have had a strong effect on their employment chances, a finding that should be approached in future research.

[28] This is the case because the comparison within the roofing sector assumes a zero effect of the MW for workers without a binding MW.

for workers falling in different wage deciles. This is feasible because we are able to identify comparable workers within the control sector.

4.5.2 Employment Effects Along the Wage Distribution

In the following approach, the treatment group corresponds to all workers of the roofing sector who are covered by the MW regulations, whereas workers in the plumbing sector, who would have been covered if they worked in the roofing sector, are considered as the control group. In particular, we identify those plumbers for whom the MW would have been binding if they worked in the roofing sector given their individual and firm characteristics. We do so by imputing the wage plumbers would receive in the roofing sector given their characteristics, thus applying the wage imputation described in Section 4.3 not only to roofers but also to plumbers and estimating the probability of being bound by the MW (P_{MW}) analogous to roofers. Appendix 4.A.5 indicates that the distribution of P_{MW} is similar across sectors, thus indicating the similarity of both sectors with respect to observable characteristics. Since the imputation of the counterfactual P_{MW} among plumbers is similar to matching individuals with a comparable treatment intensity, the necessary common support along the whole distribution seems to be given.

We then estimate the ATT for each decile $d_{it} = 1, \ldots, 10$ of the wage distribution using the following DiD framework[29]:

$$P(e_{it+1} = 1) = \kappa_d + \alpha_{g\times d} + \delta D_{it} \times d_{it} + \beta X_{it} + \gamma_{p\times d} + \nu_t + c_i + \epsilon_{it}, \tag{4.3}$$

where κ_d captures the time constant difference between workers of a different wage decile and $\alpha_{g\times d}$ captures the time constant deviation between roofers and plumbers of the same decile. Furthermore, $\gamma_{p\times d}$ allows for particular changes in each decile in the period after compared to the period before the MW introduction that are common to both group, ν_t captures year-specific effects, while the same set of covariates X as before controls for observable differences across workers. The treatment indicator D_{it} equals one for covered roofers in the period after the MW introduction (1997-2007) and zero for plumbers as well as roofers in the ex-ante period

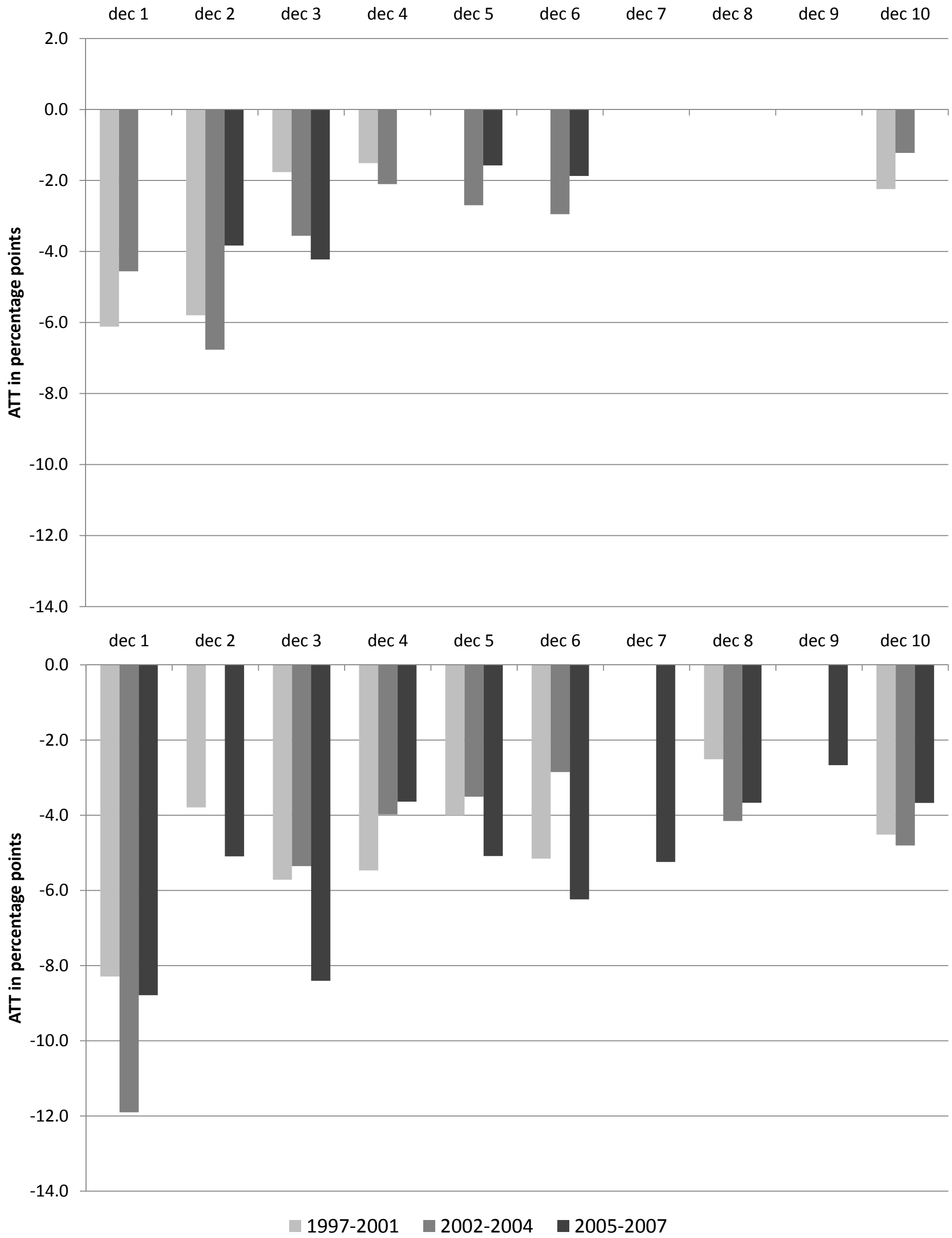
[29]The approach is related to the study of Neumark et al. (2004), who also study the MW effects throughout the wage distribution. Compared to our study, the authors exploit the regional and time variation of the MW level in the UK and look at next years' employment status along the wage distribution, defined as the distribution of initial earnings relative to the old MW.

(1995-1996). The indicator D_{it} is interacted with d_{it} so that we get an ATT for all wage deciles. Note that since the pool of workers in a particular wage decile may vary across time, we again use a fixed-effects linear probability model for the estimation, i.e. we exploit the change in a wage decile across time on the individual level for identification by including c_i in equation (4.3). The identifying assumption is that plumbers and roofers with the same set of covariates X and the same changes in the wage deciles would experience comparable changes in employment outcomes across time in the absence of the MW. In fact, this assumption is less strict than the assumption in Section 4.5.1 because we condition on the wage decile in addition to X. Moreover, we interact the model with the three sub-periods (1997-2001; 2002-2004; 2005-2007), since the effects may also differ by bite.

Figure 4.4 displays the corresponding ATEs on the probability of continued employment in the roofing sector in percentage points as long as the effect is significant at least at the 5% significance level. The results indicate that the prospects of continued employment in Eastern Germany have deteriorated due to the MW along the entire wage distribution in Eastern Germany. Moreover, note that the impact on workers with wages in the upper wage deciles are partially significant only for the latest period where the bite of the MW was strongest. For Western Germany, wage deciles that are not affected by a binding MW appear to be less affected by the MW. For workers whose wages fall in the 7th to 9th decile, no significant effects can be found at all. Still, there is some evidence for employment spillovers in line with the previous results because workers in the 3rd to 6th decile, for whom we find a decline in the chances of continued employment, are only marginally affected by a binding MW (see Appendix 4.A.6). Also, the effect seems to follow the extended bite of the MW since wage deciles 5 and 6 are only affected in the later periods with a higher MW level. The negative effect on continued employment of roofers in the 10th decile might be caused by voluntary quits of predominantly master craftsmen who leave the sample by deciding to become self-employed and to establish a single-person company whose share of all companies markedly increased during the observation period.

This additional analysis confirms that there are relevant spillovers in Eastern Germany whose temporal pattern confirms a link to the extending MW bite. For Western Germany, employment spillovers are less strong but seem to exist for workers earning wages just above the MW and for workers in the highest wage decile. These negative employment outcomes

Figure 4.4: Employment effects along the wage distribution, intersectoral comparison, by wage decile and sub-period, Western (top) and Eastern Germany (bottom), BA data 1995-2007



Note: The figures only display effects that are significant at least at the 5% significance level.

for workers with wages above the MW allow for two not necessarily competing explanations. On the one hand, the observed pattern may suggest that workers are substituted by capital and that the substitutability differs for different types of workers with the least skilled workers being easiest to substitute. On the other hand, the result pattern is compatible with negative scale effects that mostly, if not for all workers, dominate a positive substitution effect between different types of workers. Of course, both explanations may be relevant to some extent, albeit qualitative interviews with leading experts in the roofing sector (Aretz et al., 2011) suggest that the first may be the dominant explanation. In particular, the insiders doubt that the MW in the roofing sector had much of a scale effect while the relevance of technological advances such as the introduction of new roofing systems that reduce the necessary labour input have been stressed.[30]

4.6 Robustness

Interpreting the intersectoral comparisons in Section 4.5.2 as evidence for employment effects for workers who earn a wage above the MW threshold rests on several assumptions. First, there should be no indirect effect of the MW regulations in the roofing sector on the plumbing sector. If the plumbing sector provides some substitutes for roofing services, for example, a negative employment effect in the roofing sector would boost employment in the plumbing sector, thereby overestimating a negative impact of the MW. However, the lack of any evident improvement in the revenues realised by the plumbing sector relative to the roofing sector after the MW introduction puts doubt on such spillovers, see Figure 4.1. Furthermore, we find that transitions between both sectors are negligible and independent from the MW introduction. Both before and after 1997, only about 0.2% (0.1%) of all roofers (plumbers) enter the plumbing sector (roofing sector) in the next year.

Second, the assumption of common trends across sectors before and after the MW introduction must hold. In order to examine the pre-treatment trend in employment outcomes, we rely on the three years prior to the MW introduction (1994-1996). For Eastern Germany, Appendix 4.A.4 suggests a dip in employment chances in the roofing sector in 1996 that deviates from the trend in the plumbing sector. Indeed, placebo tests in row (1) and (2) in Table 4.5 confirm the common trend assumption between 1994 and 1995, while there are significant deviations between 1995 and 1996. The decline in employment outcomes in 1996 may hint at anticipation effects since employment outcomes for the last pre-MW year are measured just three months prior to the MW introduction in October 1997. Excluding observations for 1996, however, suggests even somewhat stronger negative effects than in column (2) of Table 4.4, cp. row (3) in Table

[30] The results are partly in line with Neumark et al. (2004), who find that workers whose wages are initially close to the MW (up to 1.3 times the MW) are most likely to be affected by changes in the wage floor. However, the authors find no evidence for spillovers in the upper part of the wage distribution, as opposed to our study.

4.5. If the dip in 1996 does not result from an anticipation effect, we should, however, not exclude this year, but adjust our estimates for diverging trends. A corresponding extension of the previous estimation that allows for diverging trends across sectors supports the previous findings (row (4) in Table 4.5). For Western Germany, Appendix 4.A.4 suggests that there was a dip in employment outcomes for roofers relative to plumbers in 1995. Thus, compared to plumbers, the placebo tests in Table 4.5 suggest a less favourable trend for roofers between 1994 and 1995 but a positive trend from 1995 to 1996. Estimations that allow for diverging trends across sectors, however, confirm the previous findings.

Table 4.5: Robustness checks

	LPM model	**Eastern Germany** coef[d]	N (in 1000)	**Western Germany** coef[d]	N (in 1000)
1) Placebo test: 1994-95 with plumbers	Pooled	0.2	99	-1.7***	181
2) Placebo test: 1994-96 with plumbers	Pooled	-3.2***	150	0.8**	267
3) Sample without 1996[a]	FE	-3.4***	400		
4) Trend-adjusted DiD[a]	FE	-1.8***	497	-1.2***	1,110
5) Extension with period-wise effects[a]					
ATT of $D_{it}xt_{97-01}$ in pp	FE	-3.1***		-1.3***	
ATT of $D_{it}xt_{02-04}$ in pp	FE	-3.6***	497	-1.8***	1,110
ATT of $D_{it}xt_{05-07}$ in pp	FE	-2.7***		-0.6**	
6) DiDiD[b,c]	FE	-5,6***	446	-7.2*	1,013
Share of $\hat{Y} \notin [0;1]$		6.9%		3.6%	
ATE for all covered workers		-1.1		-0.4	

[a] The coefficient should be compared to the corresponding ATT in Table 4.4, column (2).
[b] The coefficient should be compared to the corresponding ATT in Table 4.4, column (4).
[c] DiDiD estimation for workers with and without a binding MW within the roofing sector compared to the plumbing sector.
[d] Covariates are as in Table 4.4 in all specifications; standard errors are clustered at the individual level in all specifications; significance levels: * 5%, ** 1%, *** 0.1%

Adjusting for diverging trends based on the few pre-MW years, however, may not suffice if the common trends assumption fails in the long run. As a robustness check, we ran estimations that were extended by interacting the treatment indicator D_{it} to allow for a heterogeneous ATT for periods with distinct levels of a MW bite (1997-2001, 2002-2004 and 2005-2007) in order to examine the timing of the effect after the MW introduction. As shown in row (5) in Table 4.5, the impact of the MW in Eastern Germany was significantly negative in all three sub-periods, which was already suggested by Figure 4.4. In both parts of the country, the strongest impact occurred in the second period after the MW was raised to the level in Western Germany. This suggests that firms were able to bear the additional costs that were imposed by the MW during the last period, which was characterized by an economic revival of the German roofing sector.

All in all, the common trend assumption for the intersectoral comparison seems to hold as the estimation results appear to be robust and also show a plausible impact pattern across time.

However, in order to validly identify employment spillover effects within the roofing sector, it is, thirdly, important that the common trend assumption holds for the intrasectoral comparison as well.[31] To the extent that all roofers are affected by the same demand conditions, it is plausible to assume that changes in employment outcomes for workers without a binding MW between the ex-ante (1995-1996) and the ex-post situation (1997-2007) may be captured by the counterfactual change in employment outcomes for the treated roofers in the absence of the MW. However, for a period of 13 years that we cover in the estimations, diverging trends between workers with and without a binding MW in the absence of the MW may arise, e.g., due to skill-biased technological advances. We therefore capture the potentially diverging trends by using comparable workers from the plumbing sector as an additional benchmark. In particular, we use plumbers with and without a counterfactually binding MW to run a difference-in-difference-in-differences (DiDiD) estimation, where the treatment indicator D_{it} equals one for workers of the roofing sector with a binding MW after the MW introduction and zero for all other groups and time periods. If roofers with and without a binding MW had experienced different trends in their employment chances even in the absence of the MW, we can assume that plumbers with and without a counterfactually binding MW capture these trends. The corresponding DiDiD results in row (6) in Table 4.5 indicate that the employment effect in Eastern Germany continues to be negative, but somewhat smaller than in the intrasectoral comparison in Table 4.4. This may suggest that part of this negative effect is in fact due to a negative trend for workers with a binding MW relative to workers without a binding MW. Since representatives of the roofing sector repeatedly mentioned the need for catching up with technological progress in Eastern Germany (Aretz et al., 2011), this appears a plausible finding. For Western Germany, the difference between the DiD and DiDiD estimates are smaller and suggest that trends in Western Germany rather diverge in the opposite direction. More importantly, we find that the implied ATE for an average covered worker is smaller than suggested by the intersectoral comparison in the previous section.

All in all, the robustness checks confirm that employment spillovers within the roofing sector - as reported in section 4.5.2 - are likely to be the main explanation for the deviating results of

[31]Since there are only two years available prior to the MW introduction, placebo tests for the time of its introduction and trend adjusted estimations appear rather useless in order to examine the common trends assumption in the long run.

the inter- and intrasectoral comparison.

4.7 Conclusion

This paper analysed the impact of MWs in the German roofing sector on workers' chances of continued employment. For the identification of average employment outcomes, we contrasted the estimated MW impact when comparing the chances of continued employment in the roofing sector with a control sector and when comparing the chances of roofers with and without a binding MW. In addition, we estimate the causal impact of the MW for workers with and without a binding MW as well as along the entire wage distribution. We are, thus, able to also identify indirect effects of the MW on workers in the upper parts of the wage distribution for whom the MW is not binding. Our main conclusions are:

- On average, the MW in the roofing sector resulted in poorer chances of remaining employed according to both the intersectoral comparison as well as the comparison of workers with and without a binding MW within the roofing sector. This is especially true for Eastern Germany, where the MW level gave rise to a much higher share of (directly) affected workers of about 50% compared to 5% in Western Germany. Given the limited compliance with the MW regulations, the impact could even be stronger if compliance was fully enforced. Compared to the results in the construction sector by, e.g., König and Möller (2009) and Frings (2013), we therefore find stronger evidence for negative employment effects especially in Eastern Germany possibly resulting from a stronger bite of the MW.

- Estimates from the comparison of workers with and without a binding MW seem to be underestimated compared to estimates from an intersectoral comparison. If one is willing to assume that the common trend assumption holds and that the control sector is not affected by spillover effects, assumptions that are supported by some robustness checks, this deviation indicates that the MW also affects the employment chances of roofers who are not directly affected by a binding MW.

- Running an intersectoral comparison of employment chances along the entire wage distribution by exploiting the counterfactual position of workers of the control sector in the wage distribution of the roofing sector confirms this previous suspicion. The prospects of continued employment deteriorated due to the MW along the entire wage distribution in Eastern Germany. In Western Germany, spillovers are less strong but also exist for workers just above the MW level.

- The decline in employment chances among workers without a binding MW may indicate that scale effects dominate substitution effects and/or that MWs induce some capital-labour substitution. While both may be relevant to some extent, the latter may be the dominant force according to interviews that we conducted with sector insiders. In particular, they consider new roofing systems as a potential means of reducing minimum-wage induced labour costs but question a strong decline in output since the demand for roofing services appears to be rather price-inelastic.

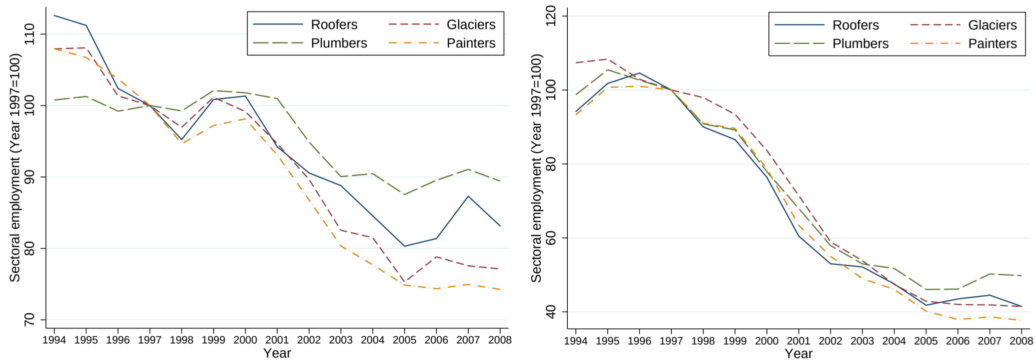
These findings on the impact of the MW regulations on the chances of continued employment should not, however, be equated with the overall MW impact on the sector's employment. In particular, the single-person companies, whose share among all companies tripled during the observation period to 23% in 2010, are not accounted for by our analysis. Furthermore, given the specific conditions of the roofing sector, e.g. the rather high level of qualification and the low labour intensity, a transferability of the results to other sectors which might be subject to MW regulations in the future has to be viewed with caution.

Despite these reservations, the presented evidence clearly highlights the need for a broader perspective on employment effects of MWs by also taking a closer look at workers who do not appear to be affected by the MW at a first glance. Moreover, our results put doubt on any attempts to identify employment effects of MWs by comparing workers with and without a binding MW within a covered sector.

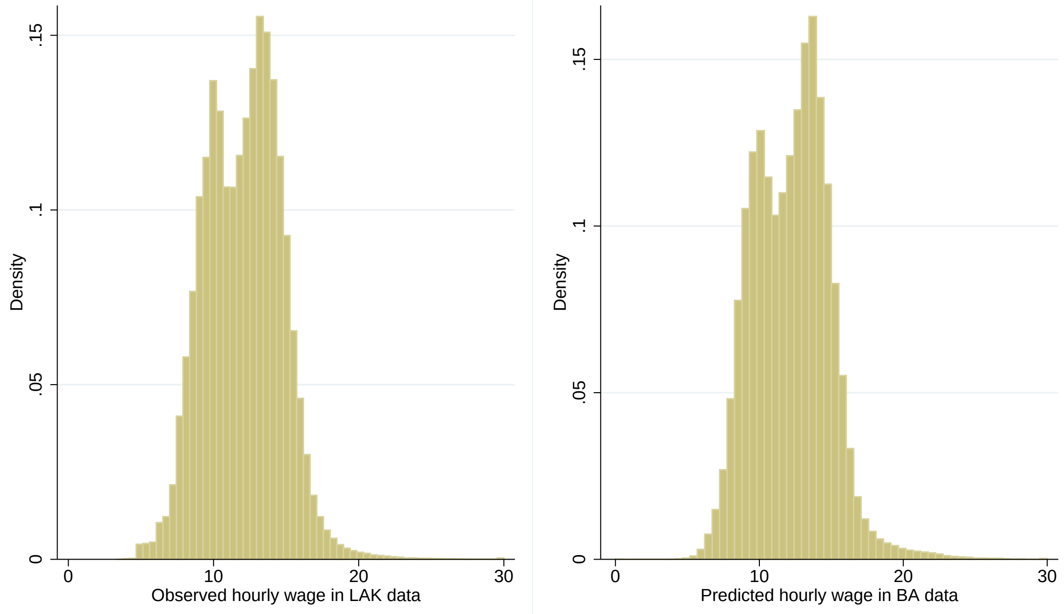
4.A Appendix

4.A.1 Evolution of the employment across sectors by region, year 1997=100, 1994-2008, BA data

(a) Western Germany (b) Eastern Germany



4.A.2 Distributions of observed LAK and predicted BA data wage



4.A.3 Explanation of the underestimation of the bite in the BA data

The compressed wage distribution which is observed in Figure 4.3 indicates that the imputation technique, which we use for implementing a hourly wage in the BA data and which assumes a normal distribution of wages, leads to a systematic underestimation of the share of workers with a binding MW in Eastern Germany. For illustration, consider two types of workers that capture the asymmetric form of the wage distribution. Type A's wage roughly corresponds to the MW, whereas type B's wage lies somewhere above the MW level. Due to the fact that there is uncertainty which individual falls into which category, the imputed distribution of the predicted mean wage always reflects a mixing distribution of these two types. As a consequence, the imputed variance for type A is an overestimation whereas type B's imputed wage variance is an underestimation of the true variance. As a result, too much probability mass for type A is above the MW and too little probability mass for type B is below the MW, resulting in an underestimation of the share of workers with a binding MW. Still, the probability of being affected by a binding MW on an individual level should be highly related to the treatment intensity, i.e. the higher this probability on an individual level is, the more likely is an individual to fall below the MW threshold and the higher is the need for increasing the wage in order to comply with the MW level.

4.A.4 Mean values of independent and dependent variables by sector prior to and after the MW introduction

Eastern Germany, BA data, 1994 - 2007

	(1)	(2)	(3)	(4)	(5)
	Roofers		Plumbers		Unconditional
	before MW	after MW	before MW	after MW	DiD
	1994 - 1996	1997 - 2007	1994 - 1996	1997 - 2007	
Dependent variable: Employed in the same sector in June of next year					
1994	0.80		0.83		
1995	0.78		0.81		
1996	0.75		0.81		
Total	**0.78**	**0.74**	**0.82**	**0.79**	**-0.01**
Individual covariates					
Age	34.45	35.80	36.04	38.48	**-1.08**
Female	0.01	0.00	0.01	0.01	**0.00**
No vocational degree	0.08	0.07	0.02	0.03	**-0.02**
Secondary education	0.81	0.88	0.89	0.93	**0.04**
Tertiary education	0.00	0.00	0.00	0.00	**0.00**
Missing educational status	0.10	0.04	0.08	0.04	**-0.02**
Skilled workers	0.73	0.78	0.86	0.87	**0.04**
Unskilled workers	0.24	0.19	0.10	0.09	**-0.04**
Master craftsman	0.03	0.03	0.03	0.04	**0.00**
Part-time <15 hours/week	0.00	0.00	0.00	0.00	**0.00**
Part-time >15 hours/week	0.00	0.00	0.00	0.01	**0.00**
Prev. work exp. in sector (in years)	1.95	9.20	1.95	7.14	**2.05**
Prev. tenure in firm (in years)	1.78	7.09	1.92	8.08	**-0.85**
Firm-level covariates					
Share of skilled workers	0.73	0.76	0.86	0.83	**0.06**
Share of unskilled workers	0.24	0.19	0.10	0.09	**-0.04**
Mean age of workforce	34.46	36.00	36.00	38.65	**-1.11**
1-5 employees	0.07	0.17	0.11	0.19	**0.01**
6-10 employees	0.16	0.25	0.17	0.19	**0.06**
11-20 employees	0.32	0.29	0.22	0.20	**0.00**
> 20 employees	0.45	0.30	0.50	0.42	**-0.07**
Mean daily gross wage in firms with > 2 workers	54.95	54.91	53.32	51.66	**1.62**
Number of observations	75,227	172,249	73,178	176,752	

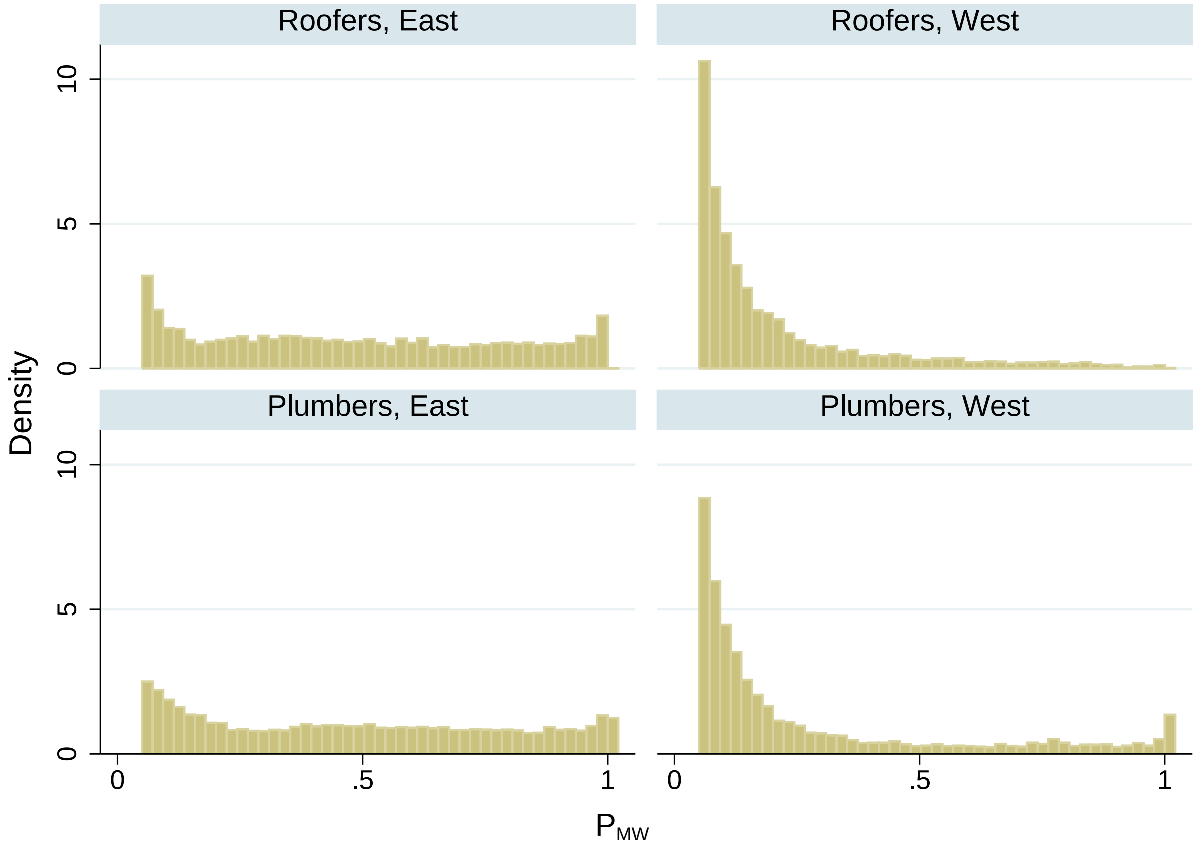
Note: The unconditional DiD is computed the following way: columns (2)-(4)-((1)-(3)).

Western Germany, BA data, 1994 - 2007

	(1)	(2)	(3)	(4)	(5)
	Roofers		Plumbers		Unconditional
	before MW	after MW	before MW	after MW	DiD
	1994 - 1996	1997 - 2007	1994 - 1996	1997 - 2007	
Dependent variable: Employed in the same sector in June of next year					
1994	0.81		0.87		
1995	0.79		0.87		
1996	0.81		0.87		
Total	**0.80**	**0.82**	**0.87**	**0.85**	**0.03**
Individual covariates					
Age	34.92	36.52	35.77	37.62	**-0.25**
Female	0.01	0.01	0.01	0.01	**0.00**
No vocational degree	0.26	0.22	0.06	0.07	**-0.04**
Secondary education	0.65	0.73	0.89	0.91	**0.06**
Tertiary education	0.00	0.00	0.00	0.00	**0.00**
Missing educational status	0.08	0.04	0.04	0.02	**-0.02**
Skilled workers	0.61	0.64	0.83	0.81	**0.04**
Unskilled workers	0.34	0.31	0.11	0.12	**-0.04**
Master craftsman	0.04	0.04	0.05	0.06	**0.00**
Part-time <15 hours/week	0.00	0.00	0.00	0.00	**0.00**
Part-time >15 hours/week	0.00	0.01	0.01	0.01	**0.00**
Prev. work exp. in sector (in years)	2.03	10.69	2.12	8.12	**2.65**
Prev. tenure in firm (in years)	1.88	8.64	2.06	9.62	**-0.80**
Firm-level covariates					
Share of skilled workers	0.61	0.62	0.83	0.77	**0.06**
Share of unskilled workers	0.34	0.30	0.11	0.12	**-0.05**
Mean age of workforce	35.00	36.88	35.77	38.07	**-0.42**
1-5 employees	0.14	0.19	0.17	0.19	**0.03**
6-10 employees	0.26	0.27	0.19	0.20	**0.01**
11-20 employees	0.31	0.30	0.22	0.22	**-0.01**
> 20 employees	0.29	0.24	0.42	0.39	**-0.02**
Mean daily gross wage in firms with > 2 workers	69.23	72.09	71.44	72.81	**1.49**
Number of observations	125,334	375,344	137,444	471,454	

Note: The unconditional DiD is computed the following way: columns (2)-(4)-((1)-(3)).

4.A.5 The probability of being affected by a binding MW, by sector and region, BA data, 1995 - 2007



4.A.6 Probability of being affected by a binding MW (P_{MW}), by year, sector, and wage decile of the wage distribution in the roofing sector, BA data, 1995-2007

Eastern Germany

Year	Sector	1	2	3	4	5	6	7	8	9	10	N
1995	Roofers	69.6	39.6	10.0	0.8	0.1	0.1	0	0	0	0	25,470
	Plumbers	74.7	36.9	8.6	0.7	0.1	0	0	0	0	0	25,208
1996	Roofers	81.0	38.9	9.4	1.1	0.1	0	0	0	0	0	26,135
	Plumbers	82.4	36.5	7.1	0.6	0	0	0	0	0	0	24,218
1997	Roofers	73.6	33.3	8.1	0.8	0.1	0	0	0	0	0	24,960
	Plumbers	75.3	31.4	5.2	0.4	0	0	0	0	0	0	23,182
1998	Roofers	63.9	29.2	8.0	1.4	0.1	0	0	0	0	0	22,489
	Plumbers	67.9	24.3	4.9	0.7	0.1	0	0	0	0	0	21,934
1999	Roofers	57.6	24.7	6.1	0.6	0	0	0	0	0	0	21,524
	Plumbers	66.4	20.6	2.5	0.4	0.1	0	0	0	0	0	21,384
2000	Roofers	73.7	52.3	28.3	9.9	1.8	0.2	0	0	0	0	18,908
	Plumbers	82.3	52.4	26.8	8.7	2.3	0.2	0.1	0	0	0	18,586
2001	Roofers	66.3	40.5	21.3	8.3	2.4	0.5	0.1	0	0	0	14,968
	Plumbers	78.9	40.3	18.7	7.7	2.3	0.5	0	0	0	0	16,164
2002	Roofers	80.3	66.6	52.5	34.9	17.1	4.0	0.5	0.1	0	0	13,076
	Plumbers	88.8	68.9	52.8	35.7	18.7	4.4	0.7	0.1	0	0	13,763
2003	Roofers	82.0	70.5	59.9	47.6	31.3	13.6	3.1	0.2	0	0	12,865
	Plumbers	88.0	70.2	60.6	47.6	30.7	13.5	3.1	0.3	0	0	12,549
2004	Roofers	85.9	75.6	67.2	56.3	42.7	26.5	10.4	1.3	0.1	0	11,681
	Plumbers	95.1	79.3	68.6	56.6	43.0	25.4	9.4	1.4	0.1	0	12,195
2005	Roofers	87.4	79.3	72.6	65.5	55.2	40.6	21.2	5.9	0.7	0	10,214
	Plumbers	86.5	76.8	70.2	63.2	53.9	40.4	20.4	5.6	0.6	0	10,806
2006	Roofers	85.7	64.0	50.8	39.8	29.8	20.8	11.7	3.6	0.3	0	10,581
	Plumbers	96.3	63.7	51.4	42.0	33.4	22.7	11.3	3.2	0.2	0	10,807
2007	Roofers	83.8	70.9	60.4	49.2	37.6	25.0	11.9	2.9	0.4	0	10,824
	Plumbers	96.8	72.0	60.3	49.3	38.1	25.2	12.8	3.8	0.4	0	11,612
N	Roofers	17,601	22,141	22,754	22,872	22,928	22,962	23,005	23,082	23,169	23,181	223,695
	Plumbers	36,831	26,062	23,183	20,713	16,990	18,468	17,675	16,391	16,578	29,517	222,408

Western Germany

Year	Sector	1	2	3	4	5	6	7	8	9	10	N
1995	Roofers	15.6	1.4	0.2	0	0	0	0	0	0	0	42,730
	Plumbers	32.6	1.2	0.2	0	0	0	0	0	0	0	46,324
1996	Roofers	20.4	2.7	0.2	0	0	0	0	0	0	0	39,239
	Plumbers	28.9	2.5	0.3	0	0	0	0	0	0	0	44,613
1997	Roofers	20.3	2.2	0.2	0	0	0	0	0	0	0	38,273
	Plumbers	33.0	2.3	0.2	0	0	0	0	0	0	0	44,087
1998	Roofers	18.4	1.8	0.1	0	0	0	0	0	0	0	36,553
	Plumbers	24.1	1.8	0.2	0	0	0	0	0	0	0	46,132
1999	Roofers	28.0	2.3	0.1	0	0	0	0	0	0	0	37,868
	Plumbers	24.6	2.0	0.1	0	0	0	0	0	0	0	46,023
2000	Roofers	41.0	7.1	0.3	0	0	0	0	0	0	0	37,959
	Plumbers	48.3	5.7	0.3	0	0	0	0	0	0	0	45,715
2001	Roofers	35.0	5.9	0.4	0	0	0	0	0	0	0	35,316
	Plumbers	53.7	5.1	0.2	0	0	0	0	0	0	0	45,329
2002	Roofers	32.8	5.1	0.3	0	0	0	0	0	0	0	33,977
	Plumbers	54.5	3.8	0.2	0	0	0	0	0	0	0	42,512
2003	Roofers	44.7	8.2	0.5	0	0	0	0	0	0	0	32,922
	Plumbers	42.7	6.4	0.2	0	0	0	0	0	0	0	39,843
2004	Roofers	58.2	14.7	0.8	0.1	0	0	0	0	0	0	31,172
	Plumbers	66.8	10.9	0.3	0	0	0	0	0	0	0	39,526
2005	Roofers	67.6	22.0	2.2	0.2	0	0	0	0	0	0	29,450
	Plumbers	67.2	20.2	1.6	0.1	0	0	0	0	0	0	37,979
2006	Roofers	69.3	23.9	2.7	0.2	0	0	0	0	0	0	29,793
	Plumbers	83.8	23.7	1.7	0.1	0	0	0	0	0	0	38,713
2007	Roofers	70.8	26	3.2	0.2	0	0	0	0	0	0	31,796
	Plumbers	84.8	25.1	2.2	0	0	0	0	0	0	0	39,298
N	Roofers	28,920	45,648	47,697	47,771	47,797	47,848	47,860	47,857	47,859	47,791	457,048
	Plumbers	45,549	58,798	52,230	51,368	52,410	47,931	43,264	38,247	45,136	121,161	556,094

5 When the Minimum Wage Bites Back: Quantile Treatment Effects of a Sectoral Minimum Wage in Germany

Single-authored

Abstract: In this study we investigate the minimum wage (MW) effects for a German sub-construction sector where the MW bites extraordinary hard by international standards. Within a quasi-experiment we estimate the Quantile Treatment Effects of the MW on the conditional and unconditional distribution of earnings. For Eastern Germany, the results indicate significant real (nominal) wage increases that ripple up to about the 0.6th quantile. However, the MW also led to declining real wages (stagnating nominal wages) among upper-decile workers, thus reducing the average pay reward for high-skilled labour in the sector. We provide evidence that a rising labour cost burden for firms together with an increased bargaining power of employers over workers still employed in the sector led to wage moderation at the upper decile, particularly among smaller East German firms. Overall this paper demonstrates how a MW geared towards the lower rank may render unexpected side effects for other workers located higher up in the wage distribution and who are mostly assumed to be unaffected by such policy interventions.

Keywords: unconditional quantile regression, minimum wages, wage effects, wage moderation, labour shortages

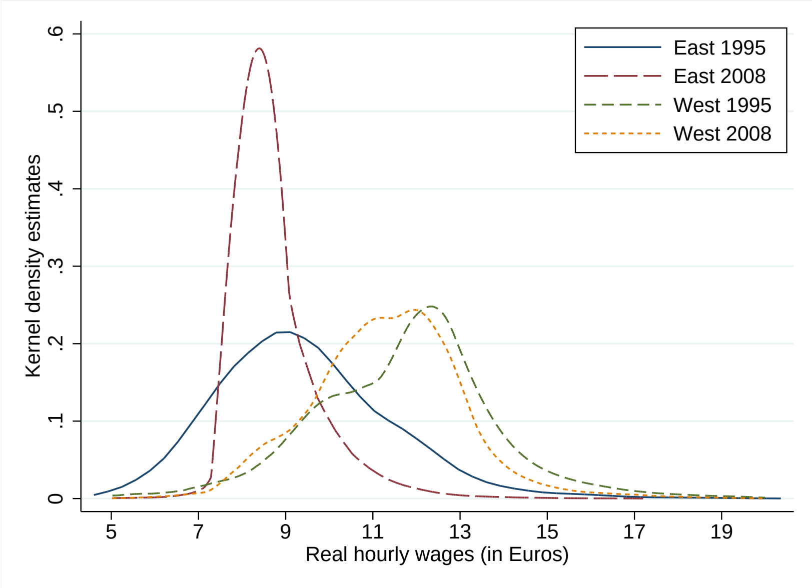
JEL-Classification: J31, J38, C18

5.1 Introduction

The way a minimum wage (MW) affects the overall distribution of earnings still remains a contested research question. Most evaluation studies focus on whether the programme improved the outcomes of particular subgroups such as (treated) low-wage workers. Such approaches primarily assume upper-decile workers to be unaffected by the policy reform and in some cases even use these workers as a counterfactual for a Difference-in-Differences type evaluation. However, these findings might not be meaningful if unexpected distributional effects appear. In fact, existing studies on MW spillovers have stated that workers with earnings above the MW may be positively affected. According to this research, wage floors create a spike in the wage distribution at the MW and boost wages of workers who earn somewhat more than the threshold. Depending on the bite, the effects then ripple up to wages at about 20% above the MW level (Neumark and Wascher, 2008). The study by Manning (2003) finds that positive spillovers can reach up to about the median worker. The conventional explanations for these findings is that firms (1) substitute low-skilled with high-skilled labour as a reaction to the change in relative input prices (Pettengill, 1981), (2) adjust their wage structure to maintain an internal wage hierarchy and hence motivation and effort among their highly paid employees (Grossman, 1983) and that (3) firms that previously paid relatively high wages to attract workers must increase wages too in order to recruit enough new employees (Manning, 2003). All these mechanisms lead to an increasing demand and thus increasing wages for workers earning a wage above the MW.

Whereas these theories are well able to explain the phenomenon of positive wage spillovers, they do not by themselves provide a complete picture of distributional aspects of MWs. For instance, recent empirical studies for the German roofing sector find a strong wage compression not only at the lower but also at the upper tail of the wage distribution, thus suggesting negative wage spillovers on high-wage earners (Aretz et al., 2012b, 2013). In particular, the descriptive findings suggest that workers in Eastern Germany with high earnings experienced deteriorating real wages in the aftermath of the policy reform (see Figure 5.1). Similar findings have been provided for the German main construction sector (Apel et al., 2012). It stands out to be tested, whether this effect is causally related to the MW or rather the result of other occurrences in the sectors. In this article we argue that in a competitive market that is facing a downward economic

Figure 5.1: Distribution of real hourly wages before and after the policy reform (LAK data)



Notes: The figure shows kernel density estimates of real hourly wages between 5 and 20 Euros based on a full sample of all roofers (see Section 5.4). Hourly wages are adjusted to prices in 1994. Bandwidths are set to 0.4.

trend and an extraordinary and increasing MW bite, firms may start compensating increasing labour costs for MW workers with wage moderation among their highest paid employees. Such wage policies become possible due to an increasing bargaining power of firms over upper-decile workers as a result of their deteriorating employment chances caused by a scale effect dominating the substitution effect. In particular, small firms that are price-takers on the market and face a less capital-intensive production technology may start exploiting such strategies to cope with the increasing labour cost burden.

The aim of this paper is to shed light on the MW effects on the distribution of earnings and wage inequality by exploiting the interesting case of the German roofing sector. For this we go beyond estimating average effects and rather focus on the heterogeneities along the wage distribution by using recent methods of quantile regression analysis. This will allow conclusions on whether the observed real wage compression in the sector is causally related to the MW policy or rather reflects an overall trend. Beyond the German case, the findings might yield important insights into possible unexpected distributional impacts of MWs (and institutions in general) on upper decile workers that have mostly been neglected in the literature so far.

This paper makes four contributions: First, we are able to study the MW effects in a context where the MW bites extraordinary hard: the roofing sector in Germany. Its MW was introduced in 1997 and was subsequently raised several times. With a Kaitz Index, i.e. the ratio of the MW level and the median wage, that is near 1 in East Germany, the bite has to be considered exceptionally high, even by international standards (Machin et al., 2003; Dolton and Bondibene, 2011). The German roofing sector thus comprises an ideal setting to study wage effects along the entire wage distribution since its bite is likely to have indirect effects on workers for whom the MW is non-binding (i.e. who have a wage above the minimum wage level). Second, we are able to exploit a quasi-experiment since, for institutional reasons, the MW was introduced only in parts of the construction sector, one of which was the roofing sector. The wage distribution of uncovered, yet comparable, sub-sectors may thus serve as a counterfactual for the earnings of roofers in the absence of the policy reform. Third, we apply an unconditional quantile regression approach recently developed by Firpo et al. (2009) that allows to investigate Quantile Treatment Effects (QTE) on the distribution of earnings. In particular, the method enables us to study the effect of the policy reform on the overall (unconditional) distribution of wages, while keeping other factors constant. To yield further insights into whether a between or within-group effect is able to explain the overall wage compression effect, we further contrast unconditional with traditional conditional quantile regression methods as proposed by Koenker and Bassett (1978) and Koenker (2005). As a final contribution, we are able to exploit two rich administrative data sets for the analysis, one of which contains a full sample of all roofers including detailed information on hours worked and the other containing a rich set of worker and firm characteristics for the treated and several untreated control sectors.

Overall, we find that the mean impact of the MW seems to miss a lot. In particular, our results suggest significant real wage increases of about 12% for lower-decile workers that ripple up to the 0.6th quantile in Eastern Germany, whereas the weaker wage effects in Western Germany (5% at the lower tail) pillar up to about the median worker. However, the estimates also reveal some unexpected side effects of the reform. According to the estimates, the MW caused a reduction in real wages by up to 5% in Eastern Germany (stagnation of nominal wages) for the highest quantiles that mostly comprise skilled and experienced workers. The wage compression effect thereby not only reflects lower entry wages, but rather indicates wage restraints among upper quantile workers, particularly within smaller firms. Contrasting conditional with unconditional quantiles further reveals that the overall wage-compression effect is solely driven by a reduction in upper tail between-group inequality, thus suggesting deteriorating returns to observable skills in the sector. The findings might explain the recent labour shortages as reported by sector insiders in separate expert interviews and/or the increase in sole-traders as reported by Kraft et al. (2012).

The structure of the paper is as follows. In Section 5.2, we give a short literature review on recent studies dealing with MW spillovers. Section 5.3 describes the German roofing sector and discusses potential control groups for a quasi-experiment. In Sections 5.4 we discuss the data base before providing descriptive evidence on the MW bite as well as on wage developments by quantile in Section 5.5. In Section 5.6, we discuss the estimation approach and provide both conditional and unconditional quantile regression estimates of the MW effect on the earnings distribution. Finally, Section 5.7 concludes.

5.2 Literature on Minimum Wage Spillovers

There are several studies that discuss how a MW affects the earnings of workers higher up in the wage distribution. Overall, the existing empirical studies suggest that MWs create a spike and boost wages for workers who earn somewhat more than the MW. The effects then fade out higher up the wage distribution, depending on the bite. Theoretically, the literature offers several explanations for these findings.

First, firms may substitute unskilled with high-skilled labour as a reaction to the change in relative input prices as suggested by Pettengill (1981). In turn, the demand for higher skilled services and therefore wages of high-skilled workers increases. The substitutability thereby decreases, the further away workers are in the skill distribution. The underlying idea is that high-skilled workers cost a lot more than low-skilled workers, but are only marginally more productive in tasks otherwise performed by low-skilled labour. Spillover effects should therefore be stronger for workers earning a wage just above the MW (and who are close substitutes), whereas workers further up the wage distribution should be unaffected.

A second mechanism has been put forward by Grossman (1983) who argues that firms adjust their wage structure to maintain an internal wage hierarchy as a result of the wage floor. In particular, they raise wages of workers earning above the MW in order to keep up their incentive schemes and maintain motivation and effort. The model leads to the same predictions as in

Pettengill (1981). Empirically, Grossman (1983) looks at the wage effects of subsequent increases in the US federal MW for workers in different occupations. She finds significant positive wage effects in the short run not only for low-paid occupations, but also for occupations that pay slightly more than the MW. The effects in the long-run are less clear in her study.

As a third explanation, Manning (2003) argues that firms that previously paid relatively high wages to attract workers must also increase wages in order to recruit enough new employees. According to his model, the effects are strongest for firms that pay just above the MW, hence spillovers will again be strongest for earnings just above the MW. For their empirical analysis, the author extends the model by Lee (1999) to include a spillover parameter for each decile. The size of the spillover effect thereby depends on the gap between the wage at a certain decile and the MW. The model predicts that the spillover effect is largest for those just affected by the MW and declines as one moves away from these wages. To identify the spillover parameter, he compares a latent and observed wage distribution. The estimates show that the maximum spillover effect is 11% for workers whose wages are near the MW. The effect then declines as one moves up the distribution and fades out at about the median wage.

An alternative explanation put forward by Falk et al. (2006) is that positive MW spillovers may reflect firms reactions to changes in the workers reservation wages and fairness perceptions. In their laboratory experiment, 91% of the workers reported reservation wages that were below the MW, whereas 59% (41%) reported an equal (higher) reservation wage after the introduction. The authors argue that workers become used to receiving a relatively high wage and develop claims such that workers think they have a right to receive higher wages and are willing to defend them. Spillover effects then arise because reservation wages of higher skilled workers depend on the actual wages of unskilled workers who earn more after a MW increase.

Finally, there is a further study that stresses the importance of wage-setting institutions in this context. For instance, Rattenhuber (2014) investigates the distributional impacts of MWs in the German construction sector and thereby distinguishes between different bargaining regimes. In particular, she finds that wage spillovers were relatively large for middle rank workers under a collective agreement compared to individually bargained contracts which, in turn, helped keep the gap to wages of the lowest paid workers stable. In line with this argument, one could argue that a moderate wage policy of unions, for instance in an economic downturn phase, might lead to a wage moderation also among higher-wage workers.

All studies discussed sofar suggest upper-decile wages to be unaffected or may even benefit if the spillovers reach far enough up the wage distribution. However, as Neumark et al. (2004) suggest, a MW may also deteriorate wages of the upper wage deciles. In particular, the authors estimate the MW effects along the entire wage distribution, distinguishing between short- and long term effects (i.e. allowing for lagged effects). The authors compare workers in US states in which the MW was raised to workers in control states where the MW stayed unchanged. The authors find elasticities of about 0.8 in the short run for worker groups earning slightly more than the MW. The elasticities are smaller for higher wage groups, but still amount to 0.15 for workers that earn between 1.5 and 2 times the MW. In the long run, the effects are much weaker and even become negative for high-wage earners. According to the authors, the finding might be driven by the scale effect resulting from higher overall labour costs that outweighs the substitution effect. The reasoning is as follows. In a simple neoclassical setting with two labour inputs, the scale effect implies a lower use of all inputs, which puts downward pressure on the demand for skilled labour (see also Neumark and Washer, 2008). On the other hand, the substitution effect implies a shift towards high-skilled labour as a reaction to the chance in relative input prices, as long as the two inputs are perfect substitutes. If however, higher costs cannot be forwarded to consumers by increasing prices or if skilled and unskilled workers are complements, the scale effect might outweigh the substitution effect, thus resulting in a lower demand for skilled labour.

This mechanism might be explaining the deteriorating employment effects among skilled workers in the German roofing sector as suggested by Aretz et al. (2013). The author investigate the employment chances of roofers along the wage distribution by comparing roofers with counterfactual outcomes of plumbers in a control group design. The authors find deteriorating employment perspectives for upper-decile workers in the sector, arguing similar to Neumark et al. (2004), that the scale effect resulting from higher overall labour costs outweighs the substitution effect. As an alternative explanation, the authors argue that the MW induced some capital-labour substitution, assuming that it did not change the relative demand for skilled and unskilled labour much. Interviews conducted with sector insiders suggest that new roofing systems provided roofing firms with possibilities to reduce MW induced labour costs in all skill-groups. In both cases the reduction in demand for higher-skilled workers might have put pressure on the salaries of these workers and thus decreased the returns to skills in the sector.

The authors however did not investigate the implications for pay rewards, as their main focus was on employment effects.

5.3 The German Roofing Sector

5.3.1 Market structure

The roofing sector is a sub-sector of the construction sector and constitutes a traditional craft that provides services including the installation of roofs on new buildings for public and private clients, repairing of roofs including energy-efficient upgrading and the installation of solar collectors. Compared to other sectors it is very capital intensive, labour costs amount to about 40% of total costs. Most firms operating on this market are relatively small and offer their services locally, thus facing a limited number of competitors[1]. There are only a few large firms with more than 100 employees that specialize in public contracts and that offer their services on a more competitive national market. A more recent trend in the sector is the increasing number of self-employed suppliers. Compared to other sectors, the craft is highly regulated as reflected by the master craftsman's diploma that is required for offering services on the market. Moreover, with a share of craftmens and skilled workers of around 70%, roofing companies operate with a relatively skilled staff. Possibilities for substituting roofing services are rather limited on this market, since they require specific skills and equipment. With competition rather driven by quality than prices, the demand for roofing services can be considered as rather inelastic.[2] Also noteworthy, according to sector insiders, technical advances such as new techniques for simplified installation and laying are very important in the roofing sector, although this is presumably mainly true for larger firms.

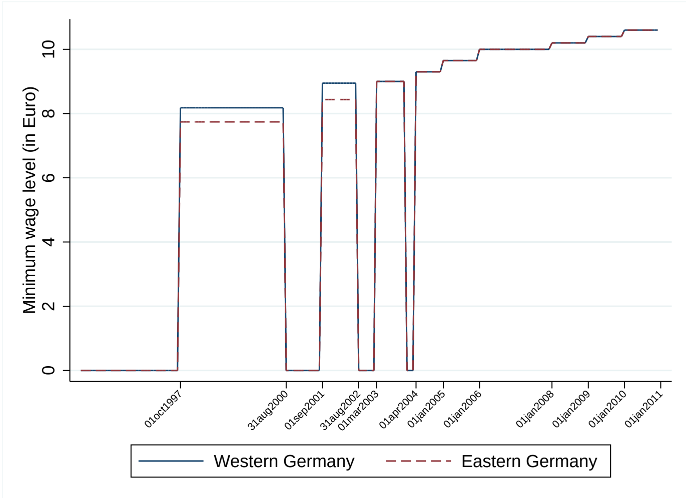
5.3.2 Minimum wage regulations

The MW in the German roofing sector was introduced in October 1997 within the Posted Workers Act (Arbeitnehmer-Entsendegesetz). The reason for the new regulation was the

[1] According to interviews, firms in the roofing sector face on average 11 competitors in their local market (Aretz et al., 2011).

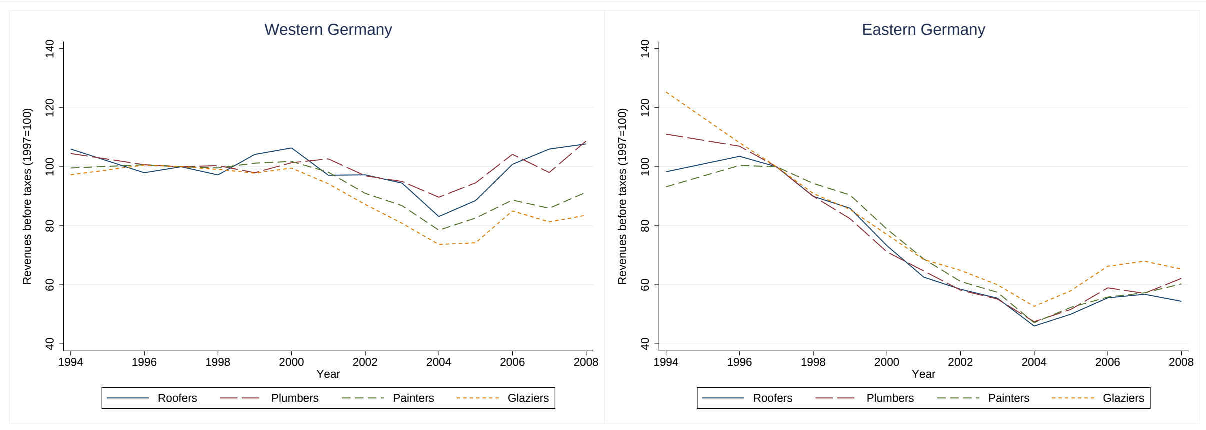
[2] Based on interviews with 250 roofers, 55 (33) East (West) German roofing firms declared price increases as their preferred strategy to account for MW induced cost increases.

Figure 5.2: Minimum wage level in the German roofing sector



European agreement on the free movement of labour that allows Eastern European firms to send workers to construction sites in other member states while paying home country wages. In order to protect the traditional craft in Germany against the increasing cost pressure from cheap Eastern European labour, the responsible trade union (Trade Union for Building-Agriculture-Environment, IG BAU) and the association of employers (National Association of Roofers, ZVDH) agreed as part of a general collective bargaining agreement on a MW of 8.2 Euros in Western and 7.7 Euros in Eastern Germany. However, since tariff agreements are negotiated on a sectoral level in Germany, not all sub-sectors implemented a MW. This provides us with the opportunity to compare quite similar sectors within a quasi-experiment. Workers in the roofing sector covered by the MW regulations comprise all blue-collar workers including minor employment. Apprentices, cleaning staff and white-collar workers are exempted from the regulations. Since 1997, the MW has been raised subsequently (compare Figure 5.2). The strongest increase thereby occurred in March 2003 for Eastern Germany, where the trade unions and employers agreed on a national MW of 9 Euros. Periods with no MW regulations are the result of tariff agreements that expired before the new regulations came into force. The interruptions were however so short, and the continuation of the MW expected, so that downward wage adjustment during this period are unlikely.

Figure 5.3: Development of revenues for roofers and selected control sectors



Notes: Revenues are taken from the German sales-tax statistics provided by the Federal Statistical Office.

5.3.3 Business cycle trends and potential control sectors

The MW regulations were introduced in a period were the entire construction sector faced a severe and long-lasting downward trend in the aftermath of the boom period in the early 90s. The depicted revenues in Figure 5.3 show that the roofing sector as well as other comparable (uncovered) sub-sectors including plumbers, glaziers and painters experienced a similar slowdown.[3] Decreasing investments in housing and industrial buildings resulted in decreasing sales and revenues that led firms to increasingly lay off workers, especially in Eastern Germany. In fact, the number of employed blue-collar workers subject to social security contributions decreased from 70,000 to 40,000 between 1994 and 2004 (see Appendix 5.A.1). After 2004, the construction sectors almost fully recovered in Western Germany, while the recovery in Eastern Germany was rather marginal. The rising importance of energy-efficient upgrading thereby helped roofing, glazier and plumber firms to fill order books again. Moreover, roofers and plumbers further profited from increasing businesses with the installation of solar panels and solar energy driven heating systems.

There are two points worth mentioning here. First, options for roofers to find a better local employment in one of the closely related sub-sectors is very unlikely in Eastern Germany. This may also explain why we find only few job transitions between the sectors (see Section 5.6). Second, it is very unlikely that roofers took advantage of the more stable West German economy, due to the low degree of residential mobility of these workers.

[3]Note that, although available in the data, we dropped structural engineers since they experienced a somewhat diverging overall trend compared to the other sectors.

Table 5.1: Various economic indicators for roofers and selected control sectors

	Roofers (1)	Plumbers (2)	Glaziers (3)	Painters (4)	Source (5)
Number of firms	10,960	13,860	3,307	8,392	A, 1996
Number of employees	87,175	94,028	16,067	51,997	A, 1996
Avg. number of employees per company	18.3	29.6	16.0	27.4	A, 1996
share of firms by revenues (in 1,000):					B, 1996
<100 DM	6.8	8.8	13.6	18.2	
100-500 DM	24.6	33.7	42.6	49,0	
$500\text{-}1,000$ DM	26.1	23.5	21.5	18,0	
$1,000\text{-}2,000$ DM	25.1	19.3	13.5	9.4	
$>2,000$ DM	17.4	14.6	8.5	5.4	
value added in € per employee	37,195	35,949	32,931	32,931	C, 2001
investments/employee (in €)	1,472	1,229	2,482	1,057	D, 2001/2002
share of labour costs (in %)	38.5	34.2	49.3	49.3	C, 2001
avg. gross daily wage/fulltime employee (in €)	66.2	68.6	66.3	67.9	A, 1996
number of companies/1 Mio. sector revenue	1.5	1.5	2.2	3	B, 1996
share of covered blue-collar workers: (in %)					A, 1996
unskilled	29.2	11.1	19.3	12.3	
skilled	66.5	83.1	73.6	82.9	
master craftmens	3.7	5	6.1	4.1	
part-time workers	0.3	0.5	0.7	0.4	

Notes: A - BA data projected to 100%; B - German sales-tax statistics of the German Federal Statistical Office (Umsatzsteuerstatistik); C - Cost Structure Survey of the German Federal Statistical (Kostenstrukturerhebung) for firms with 20-49 employees; figures for painters and glaziers are aggregated; D - Business and investment survey in the construction sector (Unternehmens- und Investitionserhebung im Baugewerbe)

Since the MW regulations were only implemented in parts of the construction sector (including the roofing sector), we are able to exploit the development of other uncovered sub-sectors as potential control groups to mimic the counterfactual development in the roofing sector. To see how comparable the other sub-sectors are with respect to their market structure, Table 5.1 shows some important market indicators for these sub-sectors. We contrast the figures of the selected industries to those of roofers for the year before the policy reform in order to test whether they might provide a potential control sector for a quasi-experiment. Overall, the comparison shows a very similar market size between roofers and plumbers in terms of firm counts and revenues. In fact, the number of companies per 1 million sector revenues, as a measure of competition, is identical. Also, value added, investments per employee as well as labour cost shares of roofing companies is more closely related to plumbing firms compared to other sectors. We thus use plumbers as a suitable control group for a Difference-in-Differences analysis (see Section 5.6).

5.4 Administrative Linked Employer-Employee Data

For the analysis we are able to exploit two administrative data sets including a full sample of all roofers provided by the Central Pay Office (Landesausgleichskasse, LAK) as well as large subsamples of all roofers, plumbers, glaziers and painters subject to social security contributions available from the Federal Labour Agency (Bundesagentur für Arbeit, BA). Both data sources are discussed in detail below.

5.4.1 LAK Data

The LAK is a public service institution of the employer association ZVDH and the trade union IG Bau in Germany. The main objective is to help insure employees against several structural disadvantages of the sector. For instance, the agency compensates roofers for earnings losses caused by bad weather, ensures a thirteenth monthly income, administrates working-time accounts and old age benefits and promotes vocational education in the sector. For these purposes, the office collects monthly information from firms on the number of actual working hours for each worker as well as their gross wages and the length of their current employment from the year 1995 onwards. Since the reporting is mandatory for firms, and may impose a penalty for non-compliance, the information is highly likely to comprise all blue-collar roofers. The information is complemented with further worker characteristics including the date of birth and sex of workers as well as an establishment identifier to calculate further firm-level characteristics. Since the data does not comprise information on education and training, we drop workers below 19 years of age that should eliminate most apprentices that are not covered by the MW regulations. Furthermore, we focus on men only, since female workers account for only a small fraction in this sector (less than 2%). Moreover, we drop observations where workers are reported to be sick, on vacation, serving in the military, and those with missing and unrealistically high (or low) wages and drop minor employment. Finally, we focus on monthly observations in June to make the data comparable to the BA data and to avoid distortions due to seasonal fluctuations during the months October to April where compensation payments by the LAK are more relevant. In total, we are able to exploit 1,055,137 June observations for 206,753 roofers across the period 1995-2010. Note that most descriptives in the following analysis are based on this very precise data set. However, the main disadvantage of the LAK data is that it is only available for the roofing sector, thus precluding the possibility of an inter-sectoral comparison. Furthermore, information on the education level and skills of workers is missing. For this reason, we exploit another data set discussed below.

5.4.2 BA Data

As a second data source we use Linked Employer-Employee Data from the Institute for Employment Research (IAB) that matches representative annual establishment survey data from the IAB Establishment Panel with personal data generated in labour administration and social security data processing. In particular, we use subsamples of roofers (75%), plumbers (30%), glaziers (75%) and painters (25%) subject to social security contributions by their employers. The data includes individual employment histories for these workers on a daily basis including several worker characteristics such as the age, sex, occupational status, gross daily wages and education of workers. The firm-level information comprises information on the workforce structure including the number of workers in certain educational groups. For the analysis, we use annual cross sections at the cut-off date June 30th. Similar to the LAK data we focus on male workers above 19 years of age and drop minor employment. As an advantage of the data set, we are able to identify and drop apprentices and white-collar workers that are not covered by the MW regulations. In total, the data set comprises 788,665 yearly observations for 171,194 roofers as well as 1,522,014 observations for 340,095 workers from uncovered control sectors for the time period 1994-2008. The main disadvantage of the BA data set is the lack of hourly wages. We thus use daily wages for our inter-sectoral comparisons. In order to compare the LAK and BA data set we further impute the wages of roofers from the LAK to the BA data to get a probabilistic MW affectiveness measure. In short, we regress hourly wages in the LAK on a set of covariates that are included in both data sets. Based on the coefficients we then estimate hourly wages in the BA data set taking into account the imprecision of the estimates. Note that we use the imputed wages for the descriptive analysis only.[4]

[4] A detailed description of the imputation procedure is available from the author upon request.

5.5 Minimum Wage Bite and the Development of Wages

Table 5.2 shows several indicators of the MW bite in June preceding the new regulations within the next year for both workers with a binding and non-binding MW. A MW is thereby defined as binding if the salary of a worker exceeds the upcoming threshold. The worker is then said to be (directly) affected. Columns (1) and (2) show the fraction of affected workers in the LAK and BA data. In order to get a first impression of how the policy reform affected earnings, Column (3) displays the individual wage gap, which we contrast to actual wage growth of binding and non-binding workers in Columns (4)-(5). The individual wage gap tells us how much wages should have increased, on average, if firms had fully complied with the new regulations. For an international comparison, Column (6) further provides the Kaitz-index that is defined as the ratio between the MW level and the median wage in the sector. Note that the indicators may slightly underestimate the bite due to the fact that hourly wages may contain overtime compensation that is not subject to the MW.[5]

The indicators show large differences between Eastern and Western Germany. For Western Germany, the descriptives show a relatively low bite. Starting with 3.8% in 1997, the share of workers with a binding MW increased moderately to 5.2% in the course of the 2008 MW hike, before dropping again slightly in the year after. The results based on the BA data perform quite similarly, showing that the imputation procedure works quite well. According to the Kaitz-Index, the bite in West Germany lies in the range of what has been found for other countries. For instance, Dolton and Bondibene (2011) compare the MW bite across 22 OECD countries and find rates between 30% and 70%. Looking at the figures for wage growth among West German workers reveals only partial compliance with the MW regulations, although improving slightly at the end of the observation period. The latter might be explained by stronger controls after 2006 that have been reported by sector insiders (Aretz et al., 2012b). Despite the lack of compliance, the figures for nominal wage growth range between 3.6% and 8.1% for affected workers, whereas the salaries of non-affected workers increased only moderately by 0.7-3.3%.

The findings for the Eastern part of the country look quite different. According to Column

[5] Overtime hours account for 6% of the working hours in June. This may lead to an estimated hourly wage that is up to 1.6% too high depending on the applied overtime compensation scheme ranging from no additional compensation to a markup of 25%. Since we do not know which scheme is applied and since the resulting imprecision appears to be rather marginal, we left the data uncorrected.

Table 5.2: Indicators of the minimum wage bite measured in June prior to the next minimum wage regulations (LAK and BA data)

		Workers with a binding minimum wage?					
		Yes				No	
New MW regulation takes effect on	MW (in Euro)	Share (in %) LAK (1)	Share (in %) BA[a] (2)	Wage gap[b] (in %) LAK (3)	Δ Wage[c] (in %) LAK (4)	Δ Wage[c] (in %) LAK (5)	Kaitz Index LAK (6)
Western Germany							
01.10.97	8.2	3.8	1.4	16.9	3.6	2.3	65
01.09.01	8.9	1.5	2.5	9.6	7.0	1.5	67
01.03.03	9.0	1.5	2.6	10.0	5.6	2.5	67
01.04.04	9.3	2.1	3.1	9.3	5.9	1.5	68
01.05.05	9.6	2.7	4.3	8.6	4.7	0.7	70
01.01.06	10.0	4.1	5.9	7.8	5.0	1.2	73
01.01.07	10.0	4.4	6.3	8.2	7.0	3.3	73
01.01.08	10.2	5.2	7.1	6.9	5.6	2.3	73
01.01.09	10.4	4.6	7.1	6.5	8.1	3.1	73
Eastern Germany							
01.10.97	7.7	13.4	8.6	12.2	6.7	-0.1	82
01.09.01	8.4	14.0	16.8	4.1	4.7	0.7	89
01.03.03	9.0	33.9	28.6	4.3	4.2	0.2	95
01.04.04	9.3	43.8	34.1	3.9	4.2	0.4	98
01.05.05	9.6	46.7	41.4	4.3	4.0	0.2	99
01.01.06	10.0	55.3	49.1	4.1	4.1	0.2	100
01.01.07	10.0	45.0	47.3	1.6	1.9	1.0	100
01.01.08	10.2	53.2	48.9	2.7	3.3	1.4	101
01.01.09	10.4	49.8	51.2	2.4	3.3	0.7	100

[a] The share is based on imputed hourly wages (see Section 5.4.2). The values thus reflect the probability of an individual to earn a wage below the new MW level.
[b] The individual wage gap is calculated as follows $wgap_{it} = (MW_{i,t+1} - w_{it})/w_{it}$.
[c] Δ wage corresponds to the actual observed percentage nominal wage change $(w_{it+1} - w_{it})/w_{it}$ between the June preceding and the June following the new MW regulation.

(1), already 13.4% of all roofers earned a wage below the 1997 wage floor. The share then increased rapidly to 34% in 2002, a few months before the national MW level was introduced. In June 2005, more than half of the workers had a wage below the 2006 MW level. In fact, the Kaitz-Index approached the value of 100, that is the median wage meanwhile equals the MW. Compared to the findings of Machin et al. (2003) for a strongly affected low-wage sector in the UK, the bite in the German roofing sector seems extraordinary hard. The mere size of affected workers in Eastern Germany also made misuse of the MW regulations harder which is reflected by the higher compliance observed. More strikingly, East German workers with salaries above the wage floor experienced almost no nominal wage growth. Even in the recovery period wages

Table 5.3: Worker characteristics for certain quantiles of the real daily wage distribution (BA-data)

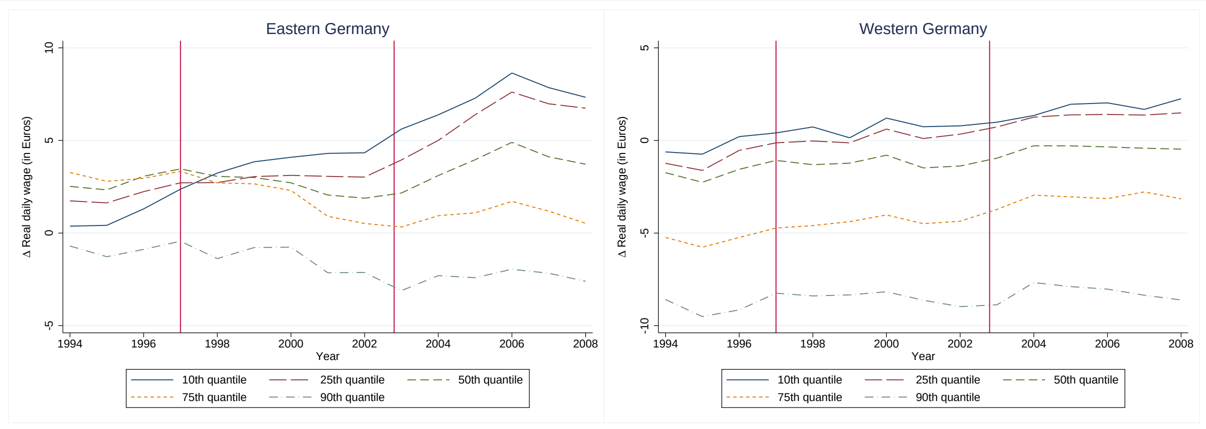
	Quantile				
	$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$
Western Germany					
real daily wage (in Euro)	52.8	61.0	68.7	74.8	81.4
yearly growth of real daily wages (in %)	-0.5	-0.3	-0.4	-0.3	-0.5
nominal daily wage (in Euro)	58.5	67.6	76.1	82.9	90.2
yearly growth of nominal daily wages (in %)	1.0	1.3	1.2	1.2	1.0
share of unskilled workers (non-technicians)	59.0	47.0	28.6	18.1	17.5
share of skilled workers (technicians)	39.2	52.0	70.4	80.2	76.5
share of master craftmen	1.1	0.8	0.9	1.6	5.9
without vocational traning degree	43.4	35.2	20.6	13.6	13.6
with vocational traning degree	56.5	64.6	79.3	86.3	86.1
with university degree	0.1	0.2	0.1	0.1	0.3
tenure in sector (in days)	783	1,114	1,521	1,863	1,963
average workforce age	32.2	33.3	36.2	38.8	40.9
average firm size	11.5	11.8	12.1	12.9	16.0
number of workers	353.7	353.9	353.9	353.8	353.7
Eastern Germany					
real daily wage (in Euro)	42.4	46.2	50.4	56.3	63.4
yearly growth of real daily wages (in %)	0.5	-0.1	-0.8	-1.3	-1.6
nominal daily wage (in Euro)	47.1	51.2	55.8	62.2	70.0
yearly growth of nominal daily wages (in %)	2.0	1.5	0.8	0.2	-0.0
share of unskilled workers (non-technicians)	27.2	22.4	18.2	14.4	15.1
share of skilled workers (technicians)	71.3	76.7	81.0	83.8	80.5
share of master craftmen	0.7	0.6	0.6	1.7	4.1
without vocational traning degree	9.2	6.0	5.7	5.1	5.5
with vocational traning degree	90.8	93.9	94.2	94.7	94.3
with university degree	0.1	0.1	0.1	0.2	0.3
tenure in sector (in days)	853	1,042	1,169	1,499	1,711
average workforce age	33.7	34.1	35.3	37.5	38.2
average firm size	13.0	13.6	14.2	16.2	18.4
number of workers	172.1	172.1	172.3	172.3	172.0

Notes: All figures shown in the table reflect average yearly values. Real wages are inflation-adjusted to prices in 1994.

increased only moderately in nominal terms. The descriptive evidence thus suggests pronounced pay restraints in this part of the country and which happened during a severe downward trend in the sector (see Section 5.3.3). As a result of this wage moderation, workers experienced deteriorating real wages resulting in a strong wage compression at the upper tail of the real hourly wage distribution in Eastern Germany (see Figure 5.1).

To get some insights into how large (or small) the wage changes were and which type of workers profited (lost), Table 5.3 displays average yearly wage changes and other worker characteristics for selected wage quantiles. According to the figures, West German roofers in all

Figure 5.4: Development of roofers real daily wages relative to plumbers (BA data)



Notes: The vertical lines represent the introduction of the MW in October 1997 and the national MW in March 2003. A positive sign reflects a higher daily wage for roofers compared to plumbers. Wages are inflation-adjusted to prices in 1994.

quantiles experienced real daily wage losses between 0.3-0.5%. The picture for East Germany reveals a much more heterogenous pattern across quantiles. Whereas workers in the lowest quantile experienced real daily wage increases by 0.5%, corresponding wages for the highest paid workers decreased by 1.6% on average. Table 5.3 gives some indication that those workers with earnings losses in the upper tail mainly comprise skilled and more experienced workers in larger firms with higher educational degrees. The mapping of skills and wages is even more pronounced in the more unequal West German distribution. Overall, the findings hint at stagnating salaries for higher skilled labour in the sector. The decreasing pay rewards for skilled labour are thereby not only driven by new roofers entering the labour market with lower entry wages, but rather reflect pay restraints among higher skilled workers that have been working in the roofing sector and for the same firm ever since.[6] Also, except for the lowest reported quantile, there was not much adjustment in terms of working hours, despite the normal business cycle fluctuations (see Appendix 5.A.3). Of course, we can not completely rule out other compensation schemes.

There might be different forces driving the observed wage compression other than the MW. First, it may simply reflect an overall trend in the sector. Second, the composition of workers may have changed. Whereas the latter can only be investigated by means of a regression analysis, the first can be explored descriptively by an inter-sectoral comparison. To provide some evidence in this regard, we use the wage developments of plumbers as a counterfactual for roofers in the

[6]For this, we restricted the data to all roofers that we observe across the entire 16-year time period (thus comprising a balanced panel) and who worked for the same firm. The descriptive findings were similar.

absence of the policy reform. Figure 5.4 shows the inter-sectoral real wage differential between roofers and plumbers over time by quantiles. The panels show that, except for the lowest earning group, wages developed quite similar between both sectors in the pre-reform period. In contrast, whereas wages of roofers in the upper quantiles decreased compared to plumbers in the post-reform period, low wage workers clearly improved in terms of daily labour income. The fact that the 10th quantile shows some anticipation behaviour will be one reason why we later conduct several placebo tests. For Western Germany, the inter-sectoral comparison reveals only marginal changes.

The decreasing pay rewards for high-skilled roofers compared to plumbers in Eastern Germany might be the result of a lower demand for these workers as suggested by (Aretz et al., 2013). The authors argue in favour of the scale effect driven by higher overall labour costs that dominate the substitution effect or some capital-labour substitutions (see Section 5.2). To test for the relevance of an increased cost burden due to the MW, Table 5.4 shows the size and change in firms' labour costs for binding and non-binding employees (as a share of total wage bill). We thereby distinguish between very small (less than 6), small (6-10), large (11-20) and very large firms (more than 20). The figures are only shown for firms with at least one MW worker. Not surprisingly, the total number of affected firms is much higher in Eastern compared to Western Germany (more than twice as high). For firms in both parts of the country, most affected firms are small firms (52% in West and 58% in East Germany), whereas the shares of larger firms with at least one MW worker are only marginal (6% and 4%). Moreover, within smaller firms, the share of affected workers is 50% and 71%. As a results, labour cost increases that firms had to bear for workers with a binding MW (as a share of total wage bill) amounts to 2.9% to 2.4% on average for small firms, while very large firms have to cope with corresponding cost increases of only 0.3% and 0.9%. Labour cost increases for workers with a non-binding MW were much lower across firm size.

The labour cost burden due to MW workers is thus much higher for smaller compared to larger firms, although the overall burden seems quite low at a first glance. However, note that these are average values before each MW hike so that these costs accumulate over time. In fact, calculating the accumulated labour cost share increase for binding MW workers shows that the annual increases accumulate to a total of 22% and 26% for smaller firms. In contrast, labour cost share increases only accumulate to 3% and 8% for larger firms, although the 8% for

Table 5.4: Average changes in labour cost shares by firm size for firms with at least one MW worker (LAK-data)

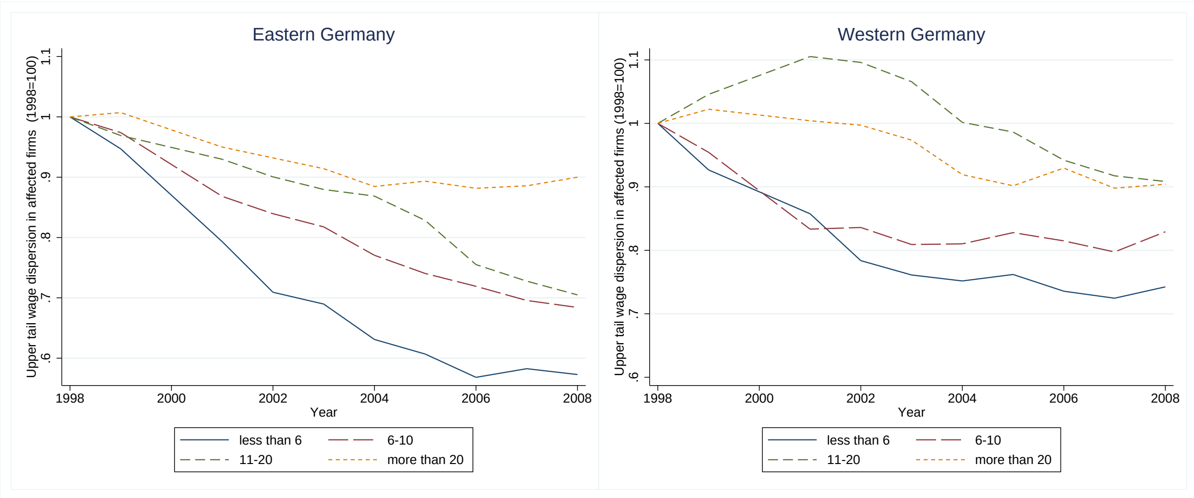
	number of employees			
	less than 6 (very small)	6-10 (small)	11-20 (large)	more than 20 (very large)
Western Germany				
number of employees	3.0	7.6	14.1	34.6
total wage bill (in Euros)	5,143	15,298	29,979	79,287
share of affected workers (in %)	49.6	19.4	12.4	7.6
labour cost share increase[a] (in %)				
for workers with binding MW	2.9	0.7	0.4	0.3
for workers with non-binding MW	1.2	1.5	1.8	1.0
accumulated labour cost share increase (binding MW)	26.0	6.2	3.8	3.0
number of firms with at least one affected worker	518.0	249.3	149.2	56.3
as a share of total firms (in %)	52.4	25.9	15.6	6.1
Eastern Germany				
number of employees	3.0	7.5	14.0	29.7
total wage bill (in Euros)	4,669	12,205	23,537	51,079
share of affected workers (in %)	68.5	50.0	43.4	37.4
labour cost share increase[a] (in %)				
for workers with binding MW	2.4	1.3	1.1	0.9
for workers with non-binding MW	0.4	0.5	0.4	0.3
accumulated labour cost share increase (binding MW)	22.0	12.0	9.8	8.4
number of firms with at least one affected worker	1,188.2	493.3	229.2	70.1
as a share of total firms (in %)	58.4	25.4	12.3	3.9

[a] Labour cost increase is calculated as follows: Δ *labour cost*$_{jt} = \sum_{i=1}^{N}(w_{ij,t+1}h_{ij,t+1} - w_{ijt}h_{ijt})$.
Notes: All values in the table are averages for the June observations before the next MW introduction including the years 1997, 2001-2008.

Eastern Germany still seems quite high. For very small firms with less than 6 employees and a total wage bill of only around 5,000 Euros (for blue-collar workers), this increase might have required some cost-reducing adjustments. This is also confirmed by interviews where almost 50% (65%) of West (East) German roofing firms expressed that they would generally need to make necessary adjustments if labour costs increased by 10% (Aretz et al., 2011). Especially smaller firms face limited options in this respect. First, they can not influence market prices much, compared to larger firms (i.e. they are price takers). Second, exploiting technological advances for capital-labour substitutions does not seem feasible either for such small firms. Therefore, pay restraints for skilled labour might have been the only option.

To test the later hypothesis more directly, Figure 5.5 shows the development of average upper decile wage dispersion within firms (measured by the standard deviation of hourly wages

Figure 5.5: Development of upper tail wage dispersion in firms with at least one MW worker (LAK data)



Notes: Panels are smoothed using a moving average. All time series are indexed to the year 1997.

among workers with earnings above the firm median wage). The panels are shown by firm size and are restricted to firms with at least one MW worker. The figures are very similar between East and West Germany and show that wage differentiation in the upper tail within firms has decreased more strongly in smaller compared to larger firms. Although not shown here, this heterogeneous development across firms can not be observed for firms with no MW worker. To conclude, we have provided several descriptive statistics suggesting that smaller firms have been more conservative compared to larger firms in increasing wages for their highly paid employees. We have also shown that this observation coincides with an increasing labour cost burden these smaller firms are facing in the course of the MW policy. In the following analysis, we will test for this causally.

5.6 Quantile Treatment Effects on the Distribution of Earnings

5.6.1 Approach

The aim of the empirical analysis is to estimate the causal impact of the MW on different parts of the wage distribution in order to investigate whether the policy reform affected the overall distribution of wages in the German roofing sector. For the identification, we are able to exploit a quasi-experiment since, for institutional reasons, the MW was introduced only in parts of the construction sector including the roofing sector. This allows us to compare the wage distribution between roofers (treatment group) and a comparable sub-sector without a MW (control group). In particular, we use the wage distribution of plumbers (see Section 5.3) to construct a counterfactual distribution for roofers in the absence of the policy reform and compare their wages before and after the MW introduction. Abstracting from the estimation procedure (conditional vs. unconditional quantile regression, see discussion below), our quantile-specific Difference-in-Differences model can be written as follows:

$$q_t^\tau = \alpha^\tau + \beta^\tau D_i + \gamma^\tau Post_t + \delta^\tau (Post_t \times D_i) + \eta^\tau \mathbf{X}_{it} + \upsilon_t^\tau + \epsilon_{it}^\tau$$

where q_t^τ is the log daily wage of workers at the τth quantile. D_i refers to the treatment variable that takes the value one for treated roofers and zero for untreated plumbers. $Post_t$ takes the value one for the post-reform period (t_1: 1998-2008) and zero for the pre-reform period (t_0: 1994-1997). $\mathbf{X}_{it}$ is a set of individual and firm level covariates including age, tenure in sector, educational attainment (6 categories), occupational status (3 categories), a part-time dummy, qualification of the workforce (3 categories) and firm size (4 categories). This rich set of covariates controls for the selection into treatment based on observable characteristics. Moreover, υ_t^τ captures any time-specific effects. By comparing four wage distributions, namely that of roofers with plumbers both before and after the policy reform, the term δ^τ gives us the Quantile Treatment Effect (QTE) of the MW introduction (and subsequent increases) on the τth quantile of the earnings distribution (for a nice graphical illustration of Differences-in-Differences with quantile regression see also Havnes and Mogstad, 2014). The interpretation thereby depends on whether one compares conditional vs. unconditional quantiles, which is discussed in detail below.

The assumptions needed for identification are that differences in wages between roofers and plumbers would have stayed the same in the absence of the policy reform (common trends). Furthermore, we need to assume that there are no indirect effects of the MW regulations in the roofing sector on the plumbing sector (no control group contamination). Concerning the first assumption, we have argued that the plumbing sector is a very comparable sector in terms of its market structure and experienced a very comparable trend in terms of important economic indicators (for details see Section 5.3). It is therefore very likely that the roofing sector would have experienced a similar wage development, had the MW regulations not been implemented in this sector, conditional on $\mathbf{X}$. Note that the inclusion of a large set of covariates as well as a set of time dummies should further ensure the assumption to hold, despite the general commonalities between the sectors. Moreover, comparing the outcome variable before the policy change shows quite a similar trend, although there are some hints for anticipation behaviour. To test for anticipation effects, we conduct several placebo tests and re-estimate the model with different specifications discussed in Section 5.6.3. The overall results however do not change much. To test for indirect effects on the control sector we further calculate transition rates of roofers into the plumbing sector. Since the share of job changes from the roofing to the plumbing sector was only 0.35% between 1994-2008, there does not seem to be much scope in terms of roofers leaving the sector and become a plumber and vice versa. Moreover, as the entire construction sector experienced a downward trend during the observation period, it is very unlikely that demand effects in the plumbing sector indirectly affected the roofing sector.

To get more insights into which period of analysis affected the wages of roofers most, we further run estimations where we interact the treatment variable with dummies for three sub-periods including 1998-2002, 2003-2005 and 2006-2008. The first period thereby reflects the period between the introduction of the MW in 1997 and the national level that was set in 2003. The years after 2003 are subdivided into two further periods to reveal further heterogeneities in recent years where the bite increased strongly. Finally, we interact the model with firm size to test whether smaller firms are driving the wage compression as suggested by the descriptives.

Quantile regression estimation. For the estimation we apply and contrast a conditional quantile regression approach proposed by Koenker and Bassett (1978) and Koenker (2005) with a unconditional quantile regression technique recently developed by Firpo et al. (2009). In case of conditional quantiles, the left-hand side of Equation 5.1 become $E[q^\tau|\mathbf{X}_{it}]$, that is we estimate the effect of the MW on the conditional distribution of wages, holding other factors constant. Note that conditional quantile regression estimates are thereby much harder to interpret as they capture the impact at certain parts of the distribution within certain groups with similar observable characteristics. This precludes any implications on the impact of a policy change on the overall observed distribution including aggregate measures such as the wage variance or the Gini coefficient. The reason for this more restrictive interpretation is that conditional and unconditional quantiles cannot be equated as in mean regression analysis, since the Law of Iterated Expectations can not simply be applied to quantiles. Put differently, conditional quantiles do not average up to their population counterpart.

We therefore mainly focus on the method proposed by Firpo et al. (2009), which allow to estimate the effect of the MW on the unconditional (marginal) distribution of wages, holding other factors constant. For this, the authors provide a technique to transform conditional to unconditional quantiles before running the regressions. In short, the method consists of two steps. In a first step, the outcome variables Y is transformed (recentered) so that it aggregates back to the overall distribution of Y. The so called Recentered Influence Function (RIF) can be expressed as the weighted probability that the outcome variable Y lies above a certain quantile. Hence, the left-hand side of Equation 5.1 becomes $RIF[q^{\tau}]$. The weighting thereby occurs through a scaling factor equal to the inverse of the density of Y evaluated at q^{τ}, which provides the right transformation since the inverse of the cumulative distribution function transforms probabilities into unconditional quantiles. In a second step, the RIF is regressed on the explanatory variables using OLS or other specifications[7].

Based on our Difference-in-Differences model, we then receive Unconditional Quantile Treatment Effects (UQTE), which we contrast to our Conditional Quantile Treatment Effects (CQTE). There are few points worth mentioning regarding the different interpretation of the two approaches. Whereas CQTE capture changes in within-group inequality (or residual wage-inequality), UQTE include both a within- and between-group inequality effect. Put differently, in CQTE analyses, the relative position of an individual in the wage distribution is determined only by the unobserved component, while in UQTE the relative position is determined by both observed and unobserved factors. As Firpo et al. (2009) note, wages might increase due to the variable of interest for low wage quantiles where both the between and within-group effects go in the same direction, but can decrease wages for high wage quantiles where the between- and within group effects go in the opposite directions. To demonstrate the differences, the authors investigate the impact of union coverage and conclude that the detected deviations between their conditional and unconditional quantile regressions may be driven by the fact that the union wage gap generally declines as a function of the (observable) skill level. In particular, we will compare both

[7]Firpo et al. (2009) also compare their results to other specifications including a Logit and nonparametric specification. Since the results do not seem to differ much between the RIF-OLS and the alternative specifications, we stick to the first.

UQTE to CQTE to derive implications on MW induced changes in the returns to observable (unobservable) characteristics such as the education and experience of workers.

5.6.2 Results

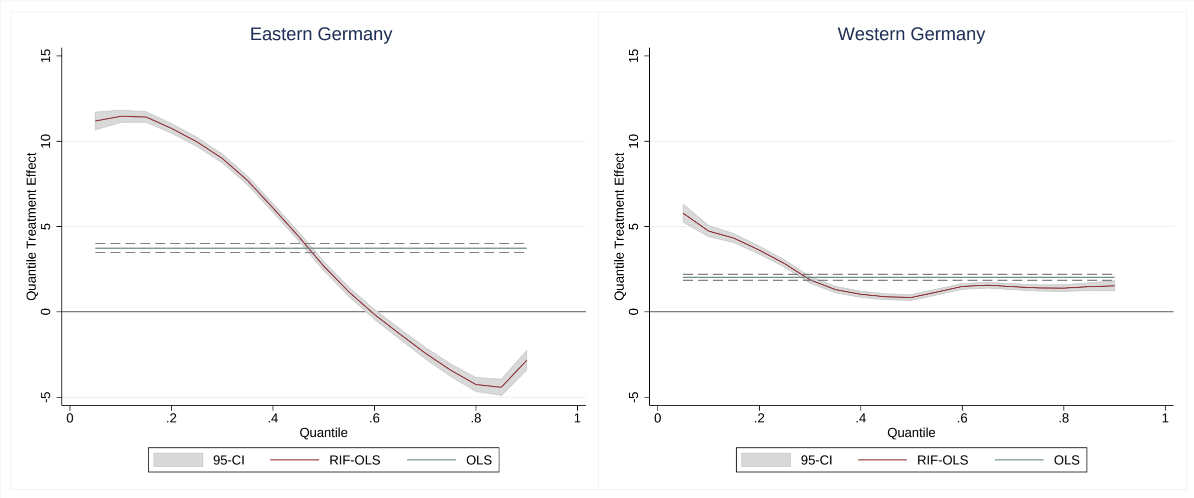
Table 5.5 displays the coefficients for quantiles τ=0.1, 0.25, 0.5, 0.75, 0.9 (Columns 2-6) as well as two inequality measures including the wage variance and Gini coefficient (Columns 7-8). As a benchmark, we further add the coefficients of a simple OLS model in Column (1). The results are shown for the basic model as well as interacted with the three sub-periods and with firm size in two separate models. The UQTE of the basic specification are also shown graphically for 19 different quantiles (from the 5th to the 95th) in Figure 5.6 including its 95% confidence interval as well as basic OLS estimates (represented by the horizontal line). The figure already reveals a large degree of heterogeneity in the MW effect across quantiles as indicated by the deviations of the RIF-OLS coefficient plots from the OLS coefficient lines.

In particular, the coefficients in Table 5.5 for Western Germany across the entire period indicate that workers at the 0.1th quantile experienced MW induced real wage increases of 5%. The effects then decrease for the 0.25th quantile to 3% before fading out at about the median worker. For higher earning groups, the wage effects are only marginal. Despite the reduction in lower tail inequality, overall wage dispersion did not change much in this part of the country, although the Gini coefficient shows a significant but small decrease.

In contrast, the corresponding figures for Eastern Germany are much more pronounced in terms of size and heterogeneity. In particular, real wages of low-wage workers were bumped up by 12%, on average, as a results of the MW. The positive wage effects thereby extend to about the 0.6th quantile (compare Figure 5.6). Also noteworthy, the MW increased the location of the wage distribution (positive effect on the median). However, despite the location shift, East German roofers with high wages experienced real wage decreases of more than 3%. As a result of the wage-compression effect at both ends of the distribution, overall wage inequality among East German roofers decreased, as shown by the significant negative effects for both wage variance and Gini coefficient. This significant wage compression effect exactly mirrors the descriptive finding in Figure 5.1.

Looking at the QTE interacted with the sub-period dummies (corresponding quantile plots

Figure 5.6: Unconditional Quantile Treatment Effects of the minimum wage



Notes: This figure shows OLS and QTE estimates including their 95% confidence intervals based on bootstrapping with 100 replications. The effects can be interpreted in percentage changes.

are shown in Appendix 5.A.5) show that the effects gradually increased with the bite. The strongest impact is thereby estimated for the last sub-period where earnings of low-wage workers were pushed up by 22% in the case of Eastern Germany. In contrast, corresponding workers at quantiles 0.75 and 0.9 faced real wage losses of 5% and 3%. The effects seem particularly strong compared to other MW studies that have looked at MW wage spillovers. For instance, estimates for a low-wage sector in the UK by Manning (2003) amount to 11% in the neighbourhood of the MW, then fading out at about the median wage.

Our estimates further indicate that firm size is important. In particular, our findings suggest positive wage effects between 2% and 9% among West German firms with more than 10 employees, irrespective of their wage position. However, in smaller firms with 6-10 employees only low-wage workers experienced real wage growth. More strikingly, workers in firms with less than 6 employees experienced zero to negative real wage growth in Western Germany. In particular, whereas real wages of the lowest quantile remained unchanged, the upper quantile workers experienced wage decreases between 1.5% and 2%. Obviously, increasing labour costs for MW workers led smaller firms to restrain wages among better paid employees. This interesting phenomenon was particularly strong in Eastern Germany and affected employees in firms with up to 20 employees. Here, real wage increases for the lowest quantile range between 11% to 13%, whereas real wage losses among the highest quantiles amounted to 4.5% and 6.5%. Even in firms with more than 20 employees, real wages among better paid workers only increased by

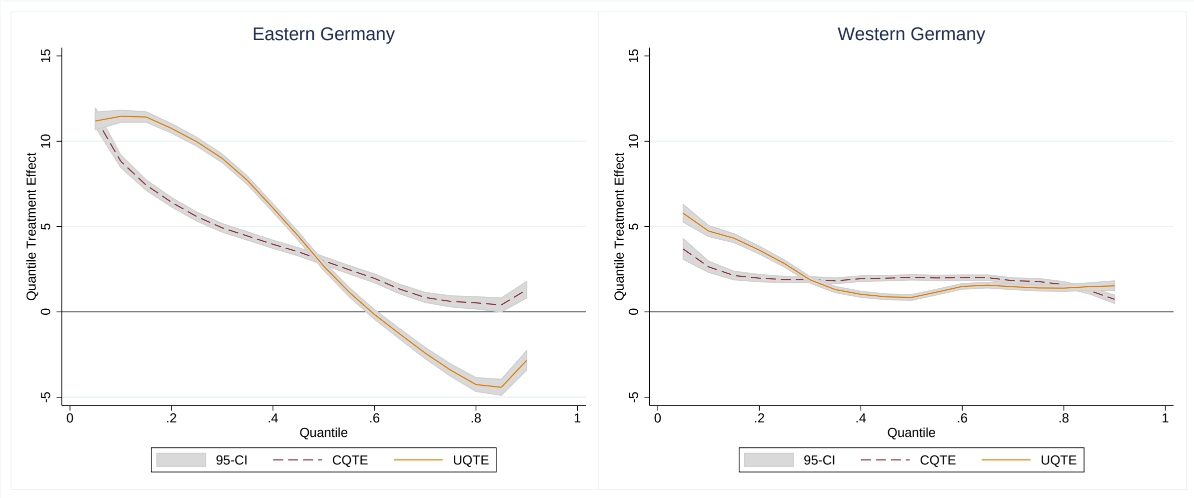
Table 5.5: Unconditional Quantile Treatment Effects of the minimum wage

Dependent variable: log real daily wages	OLS	Quantile Regression Estimates					Inequality	
		$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$	Variance	Gini
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Western Germany (N=1,055,380)								
Basic model								
1998-2008	2.04***	5.08***	2.83***	0.79***	1.36***	1.74***	-0.20	-0.14***
	(22.70)	(29.12)	(26.91)	(10.09)	(15.74)	(10.35)	(-1.81)	(-9.27)
Interacted with sub-periods								
1998-2002	1.49***	3.53***	2.20***	0.69***	0.96***	1.49***	0.13	-0.06***
	(14.46)	(17.98)	(18.53)	(7.79)	(9.76)	(7.76)	(0.99)	(-3.39)
2003-2005	2.68***	6.05***	3.57***	1.22***	1.95***	2.26***	-0.20	-0.14***
	(22.16)	(24.40)	(24.24)	(11.24)	(16.86)	(10.31)	(-1.30)	(-6.91)
2006-2008	2.53***	7.42***	3.41***	0.55***	1.57***	1.73***	-0.92***	-0.30***
	(19.80)	(27.22)	(21.77)	(4.94)	(13.80)	(7.88)	(-5.80)	(-14.67)
Interacted with firm size								
less then 6 employees	-2.19***	0.11	-0.60***	-2.19***	-1.45***	-2.02***	0.94***	0.09***
	(-19.13)	(0.46)	(-4.54)	(-23.32)	(-15.10)	(-11.23)	(7.04)	(5.30)
6-10 employees	1.70***	5.79***	3.19***	0.62***	0.19	0.16	-0.66***	-0.27***
	(16.43)	(28.41)	(25.55)	(6.63)	(1.90)	(0.88)	(-4.96)	(-15.35)
11-20 employees	3.66***	7.02***	4.24***	2.19***	2.23***	2.60***	-0.69***	-0.25***
	(34.72)	(34.54)	(33.41)	(22.61)	(20.99)	(13.20)	(-4.99)	(-14.09)
more than 20 employees	6.57***	8.83***	5.36***	3.50***	6.21***	8.14***	-0.61***	-0.13***
	(56.28)	(40.01)	(37.48)	(31.78)	(48.64)	(32.74)	(-3.79)	(-6.04)
Eastern Germany (N=451,245)								
Basic model								
1998-2008	3.74***	12.28***	9.44***	2.47***	-3.44***	-2.92***	-1.42***	-0.75***
	(27.34)	(62.93)	(75.72)	(20.49)	(-18.61)	(-9.89)	(-7.08)	(-35.66)
Interacted with sub-periods								
1998-2002	2.12***	8.38***	5.31***	1.47***	-2.33***	-2.79***	-0.67**	-0.46***
	(13.80)	(38.89)	(37.32)	(10.79)	(-11.19)	(-8.39)	(-2.95)	(-19.35)
2003-2005	4.81***	14.38***	13.33***	3.01***	-4.69***	-3.53***	-1.96***	-0.96***
	(25.92)	(48.76)	(71.29)	(16.29)	(-18.12)	(-9.23)	(-6.36)	(-29.71)
2006-2008	7.43***	21.75***	17.56***	4.88***	-5.37***	-2.60***	-3.09***	-1.41***
	(37.18)	(64.13)	(89.00)	(25.27)	(-21.07)	(-7.01)	(-9.59)	(-41.36)
Interacted with firm size								
less then 6 employees	0.61***	10.83***	7.96***	-0.37*	-7.43***	-6.49***	-1.29***	-0.79***
	(3.64)	(44.91)	(50.68)	(-2.43)	(-34.69)	(-20.22)	(-5.18)	(-29.99)
6-10 employees	2.62***	12.59***	9.22***	1.45***	-6.02***	-5.83***	-2.15***	-0.93***
	(17.01)	(58.34)	(62.56)	(9.64)	(-27.65)	(-17.91)	(-8.75)	(-36.00)
11-20 employees	3.60***	12.44***	9.41***	2.57***	-3.96***	-4.61***	-1.87***	-0.85***
	(22.69)	(58.23)	(63.42)	(16.66)	(-17.27)	(-13.45)	(-7.37)	(-31.81)
more than 20 employees	6.08***	11.44***	9.53***	4.79***	1.86***	2.43***	-0.41	-0.46***
	(34.20)	(51.73)	(62.21)	(29.58)	(7.25)	(5.96)	(-1.53)	(-16.11)

Notes: t-statistics in parenthesis. Significance levels: * 5%, ** 1%, *** 0.1%. Robust standard errors for OLS estimates. Standard errors of RIF regressions are bootstrapped with 50 replications. All models include a set of individual-level covariates including age, tenure, educational attainment (6 categories), occupational status (3 categories), part-time dummy as well as firm characteristics including qualification of workforce (3 categories) and company size (4 categories) and year dummies.

2% to 2.5% compared to 11.5% wage growth for the lowest earning group. This pattern exactly mirrors the descriptive observation that smaller firms show a large decline in their (upper-decile) wage dispersion.

Figure 5.7: Unconditional vs. Conditional Quantile Treatment Effects of the MW



Notes: This figure shows CQTE and UQTE estimates including their 95% confidence intervals based on bootstrapping with 100 replications. The effects can be interpreted in percentage changes.

To conclude, the quantile regression estimates suggest that policy makers seem to have met their goals of improving the earnings of the working poor and reducing overall inequality in the sector. However, the estimates also reveal MW induced reductions in real wages for upper-decile workers in Eastern Germany. Note that these wage reductions are not only driven by new workers entering the labour market, but rather reflect real wage decreases of workers already working in the sector and within the same firm (for corresponding tests see next Section). This interesting stylized fact might seem puzzling at the first glance. However, as noted in Section 5.5, the MW in the German roofing sector bit particularly hard in the eastern part of the economy in the sense that a higher share of workers were affected overall and within firms. Since this has strongly increased the labour cost burden of smaller firms with a share of MW workers over 70%, it may have led to wage constraints among higher-wage workers. The restrictive wage policy of firms thereby only become possible due to an increasing number of human-capital workers queuing for jobs, which has increased the bargaining power of employers over still employed workers in the sector. Nevertheless, this is striking evidence since it may render potential labour shortages in the future if firms face increasing difficulties to adjust costs otherwise. In fact, according to sector insiders and an evaluation of vacancies and job searchers (compared to plumbers) shows that the shortage of highly skilled workers has become a major challenge in the roofing sector (Aretz et al., 2011).

Unconditional versus Conditional Quantile Treatment Effects. The RIF-Regressions in the previous section reveal a substantial MW induced wage compression effect at both tails of the East German wage distribution. We now contrast these results to estimates from conditional quantile regressions in order to shed light on the forces driving the wage compression, in particular at the upper part.

Recall that CQTE measure the MW effects on the conditional wage distribution, that is changes in the wage distribution within groups of individuals with similar characteristics. Figure 5.7 plots both the CQTE and, for comparable reasons, UQTE. The detailed regression results for the CQTE are shown in Appendix 5.A.6. Looking at the figure for Eastern Germany first, shows that conditional real wages increased for the lowest quantiles but remained unchanged for upper decile workers (compare dashed line). The finding indicates that within-group inequality has decreased at the lower rank, that is for workers with similar characteristics those at the lower part of the distribution (those with unfavourable unobserved characteristics) now earn more. It may however also be the result of a selection effect caused lay-offs of low-wage workers with poor unobserved characteristics as found by Aretz et al. (2013). In contrast, for high wage earners, within-group inequality has not changed, meaning that workers with similar characteristics still earn similar relative wages. For West Germany, we find a much smaller increase in within-group inequality at the lower tail.

Comparing UQTE (solid line) to CQTE in East Germany reveals interesting deviations. In particular, unconditional estimates are much higher for workers below the median and much lower for workers with earnings above the median. Put differently, whereas CQTE suggest that real wages did not change for upper quantile workers, UQTE suggest negative real wage changes at the upper tail as a result of the MW. The fact that UQTE are higher (lower) at lower (higher) quantiles indicate that MW effects decline as a function of (observable) skills (see also discussion in Section 5.6.1). The wage compression effect at the upper tail is thus driven solely by the between-group effect indicating that wage inequality decreased between (rather than within) groups with different observable characteristics at the upper tail. The results therefore hint at MW induced decreases in the average pay-rewards to observable skills among higher-wage workers including education and experience in the sector.

Overall, the comparison demonstrates that conditional quantile regression fails to account for the observed wage changes, whereas unconditional estimates exactly mirror the compression of wages both at the top and bottom of the real daily wage distribution. It also confirms the hypothesis that the MW has reduced the returns to observable skills, so that the career incentives have decreased for higher educated and experienced workers in the sector.

5.6.3 Robustness

To test whether our UQTE in Table 5.5 are contaminated by anticipation behaviour before the policy change, we conduct several placebo tests shown in Table 5.6. For this, we restrict the sample to the pre-reform years (1) 1994-1997, (2) 1994-1996 and (3) 1994-1995 and assume each last year to be the post-reform period. If there were no anticipation effects, we would observe QTE of zero at each quantile. The results suggest positive anticipation effects in 1997 (Placebo I) and to a lesser degree in 1996 (Placebo II), meaning that firms obviously started to adjust wages upwards prior to the MW during these years, thus downward biasing the estimates. We do not find any such effects for the year 1995 (Placebo III). To make sure the results are not driven by any of these anticipation effects, we re-estimate the basic model in two versions: In a first version, we declare the 1997 to the post-reform period (Robustness I) since there are only 3 months between the data point in June 30th and the MW introduction in October 1997 of the same year. In a second version, we drop all observations of the years 1996 and 1997 (Robustness II). The slightly higher coefficients of the latter models suggest that the basic model might be underestimating the wage effects. Overall, the estimates however do not change much suggesting that our estimates are quite robust

We have argued throughout the test, that the MW caused a compression at the top of the East German real wage distribution. This compression effect might thereby be the results of two different forces. First, it may reflect new workers entering the labour market with lower entry wages compared to existing workers in the firms they start working for. Second, the real wage reductions may be the result of wage restraints among the existing workforce that already was employed in the firm before the policy reform. To test this, we restrict the sample to workers who we observe in every year in the sample, that is we create a balanced sample (Robustness III). Whereas the effects decrease strongly in Western Germany, the effects stay almost unchanged for East German roofers, which implies that our estimates are not driven by workers entering the sample. We then further restrict the balanced sample to all workers that were employed in

Table 5.6: Placebo tests and robustness checks for estimations in Table 5.5

Dependent variable: log real daily wages									
		OLS	Quantile Regression Estimates					Inequality	
Post-reform period	N		$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$	Variance	Gini
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Western Germany									
Basic model (for comparison)									
1998-2008	1,055,380	2.04***	5.08***	2.83***	0.79***	1.36***	1.74***	-0.20	-0.14***
		(22.70)	(29.12)	(26.91)	(10.09)	(15.74)	(10.35)	(-1.81)	(-9.27)
Placebo tests									
I. 1996-1997	323,402	1.51***	3.11***	2.45***	1.46***	0.62***	0.89**	0.23	-0.10***
		(10.32)	(12.25)	(13.51)	(12.16)	(4.17)	(3.28)	(1.07)	(-4.19)
II. 1995-1996	245,290	0.58**	1.92***	0.98***	0.43**	-0.38*	-0.49	-0.60*	-0.14***
		(3.29)	(6.27)	(4.39)	(2.93)	(-2.10)	(-1.47)	(-2.29)	(-5.04)
III. 1995	166,419	0.21	1.21***	0.24	-0.18	-0.67**	-0.69	-0.92**	-0.14***
		(1.02)	(3.54)	(0.95)	(-1.06)	(-3.17)	(-1.78)	(-3.10)	(-4.14)
Robustness									
I. 1997-2008	1,055,380	2.18***	5.15***	3.07***	1.05***	1.44***	2.14***	0.01	-0.12***
		(22.30)	(27.38)	(26.76)	(12.25)	(15.14)	(11.44)	(0.06)	(-7.34)
II. without years	898,397	2.52***	6.03***	3.75***	1.33***	1.59***	2.26***	0.05	-0.15***
1996 and 1997		(21.93)	(26.94)	(27.71)	(13.05)	(14.19)	(10.10)	(0.38)	(-7.83)
III. only stayers	196,193	0.43**	0.46	0.19	1.05***	0.40*	-0.30	-0.31*	-0.06**
		(2.91)	(1.75)	(1.09)	(7.93)	(2.12)	(-0.81)	(-2.41)	(-2.83)
IV. only stayers	158,894	0.47**	0.84**	0.35	1.02***	0.13	-0.76	-0.41**	-0.09***
in same firm		(2.80)	(2.85)	(1.79)	(6.45)	(0.62)	(-1.80)	(-3.19)	(-3.84)
V. nominal daily wage	1,055,380	2.04***	4.05***	3.32***	4.67***	1.36***	-3.31***	-0.70***	-0.28***
		(22.70)	(21.02)	(28.75)	(59.36)	(16.77)	(-22.77)	(-6.12)	(-19.83)
Eastern Germany									
Basic model (for comparison)									
1998-2008	451,245	3.74***	12.28***	9.44***	2.47***	-3.44***	-2.92***	-1.42***	-0.75***
		(27.34)	(62.93)	(75.72)	(20.49)	(-18.61)	(-9.89)	(-7.08)	(-35.66)
Placebo tests									
I. 1996-1997	174,680	2.70***	4.96***	4.15***	3.07***	0.23	0.04	-0.56	-0.23***
		(12.08)	(15.25)	(17.75)	(13.85)	(0.77)	(0.09)	(-1.63)	(-6.50)
II. 1995-1996	131,340	0.81**	2.17***	1.80***	1.51***	-0.73	0.04	-0.08	-0.07
		(2.84)	(5.24)	(5.92)	(5.45)	(-1.90)	(0.09)	(-0.17)	(-1.65)
III. 1995	86,755	-0.07	0.69	0.70	0.52	-1.20**	-0.75	0.19	-0.01
		(-0.20)	(1.45)	(1.95)	(1.63)	(-2.72)	(-1.26)	(0.45)	(-0.20)
Robustness									
I. 1997-2008	451,245	4.07***	12.18***	9.37***	3.08***	-2.70***	-2.94***	-1.31***	-0.73***
		(26.89)	(58.42)	(70.64)	(23.88)	(-13.28)	(-8.86)	(-6.12)	(-32.02)
II. without years	363,320	4.64***	13.70***	10.57***	3.87***	-2.83***	-4.26***	-1.50***	-0.82***
1996 and 1997		(25.29)	(55.50)	(68.91)	(26.76)	(-11.49)	(-10.28)	(-6.20)	(-30.73)
III. only stayers	41,451	3.29***	11.76***	7.75***	0.85	-5.36***	-2.80**	-1.27***	-0.60***
		(7.78)	(26.33)	(21.57)	(1.86)	(-8.04)	(-2.74)	(-4.76)	(-11.33)
IV. only stayers	31,115	3.41***	11.23***	8.99***	1.84**	-6.67***	-1.71	-1.98***	-0.70***
in same firm		(6.36)	(23.23)	(21.39)	(3.18)	(-7.55)	(-1.22)	(-5.55)	(-10.33)
V. nominal daily wage	1,055,380	3.74***	5.27***	5.76***	5.38***	0.73***	0.17	-0.53**	-0.30***
		(27.34)	(23.67)	(40.28)	(41.26)	(4.29)	(0.65)	(-2.60)	(-14.64)

Notes: robust t-statistics in parenthesis. Significance levels: * 5%, ** 1%, *** 0.1%. Standard errors are bootstrapped with 50 replications. All models include a set of individual-level covariates including age, tenure, educational attainment (6 categories), occupational status (3 categories), part-time dummy as well as firm characteristics including qualification of workforce (3 categories) and company size (4 categories) and year dummies.

the same firm during the entire observation period (Robustness IV). Again the results are quite stable for Eastern German workers, thus suggesting that the effect is also not driven by workers that lost their job and returned to work with a lower salary in a different firm. Overall, we can conclude from these robustness checks that the wage-compression effect in Eastern Germany actually reflects wage restraints on long-standing employees in the roofing sector that worked for the same firm ever since.

Although we conducted our analysis for real daily wages, as this is certainly the more relevant variable from a welfare perspective, we also check how our results change when using nominal daily wages as a dependent variable. In Eastern Germany, nominal wage increased for the lowest quantiles less than half as much compared to real wages. The effects ripple up to above the median wage as with real wages. For the top part of the nominal wage distribution we find zero and insignificant wage change. The results confirm that firms have been very conservative in their wage-setting, which, as prices increased, led to deteriorating real wages. The results for Western Germany using nominal wages are somewhat surprising. Whereas wage increases of lower-quantile wages were again somewhat lower in nominal terms, upper-quantile workers in Western Germany experienced negative nominal wage growth.

5.7 Conclusion

This study is motivated by recent descriptive findings for a German sub-construction sector that experienced a strong wage compression not only at the bottom but also at the top of the wage distribution in the aftermath of a MW introduction. In order to investigate whether the observed wage decreases for high-wage earners are driven by the policy reform, we estimate the MW effects on earnings for each quantile of the distribution. For the identification, we apply recently developed Unconditional Quantile Regressions suggested by Firpo et al. (2009) within a quasi-experiment. In particular, we are able to compare the wage distribution of the treated sector (roofers) with the distribution of an untreated control sector (plumbers) before and after the institutional change to derive Unconditional Quantile Treatment Effects. We further contrast the results to Conditional Quantile Treatment Effects in order to shed light on the wage compression effect at the upper tail of the distribution.

The results reveal large heterogeneities along the distribution, suggesting that the mean impact misses a lot. In particular, we find significant real daily wage increases of about 12% for lower-decile workers that ripple up to the 60-percent quantile in Eastern Germany, whereas the weaker wage effects in Western Germany (5% at the lower tail) pillar up to about the median worker. Here the policy seems to have met its goal of improving the earnings of the working poor and reducing overall wage inequality. However, the estimates also reveal some unexpected side effects of the reform. According to our estimates, the MW caused a reduction in real daily wages of about 5% in Eastern Germany for the highest quantiles (stagnating nominal wages) that mostly comprise skilled and experienced workers. A comparison between unconditional with conditional quantile regression estimates further shows that the wage compression effect at the upper tail is thereby solely driven by a between-group effect, which suggests that the returns to observable skills have decreased among skilled workers as a result of the MW. Estimates by firm size suggest that the negative spillovers have mainly been taking place within smaller firms. The decline in pay rewards for skilled workers thereby not only reflects lower entry wages, but rather indicates wage restraints among experienced workers that have been working for the same firm ever since.

This interesting stylized fact for Eastern Germany is the result of several occurrences in the sector. First of all, the increasing bite of the MW strongly raised the labour cost burden, particularly among smaller firms with a share of MW workers over 70%. Our calculations show that the accumulated labour cost share for smaller firms increased strongly during the observation period. This increasing labour cost burden went along with a severe downward trend in the entire construction sector since 1995. Especially smaller firms with a limited influence on market prices (price-takes) and less possibilities for substituting labour by capital have limited the scope for wage increases among their skilled employees. However, wage restraints among highly paid workers then only became possible due to deteriorating employment chances also among upper-decile workers as suggested by the complementary study of Aretz et al. (2013) and which may have increased the bargaining power of firms over still employed workers. As a result, wage differentiation and thus incentives for human capital investments have been shrinking in the sector. The findings might explain rising labour shortages that firms are facing recently as reported by separate interview with roofing firms. The less attractive working conditions for the more qualified workers may also be a reason for the MW induced rise in sole traders in the sector as suggested by Kraft et al. (2012). According to the authors, some of these qualified workers may have opted to start their own business to increase their earnings.

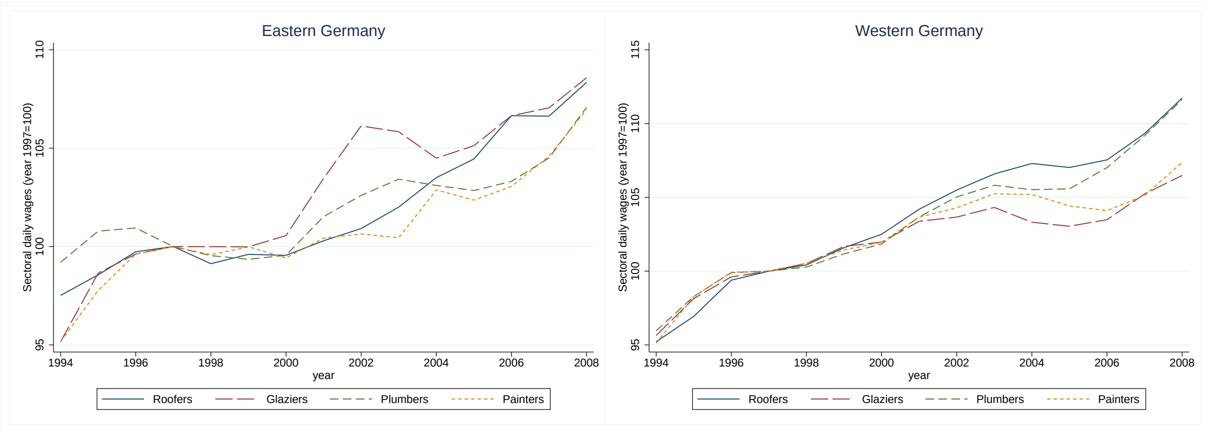
Although the investigations were conducted for the German roofing sector only, a MW might induce similar effects in other comparable sectors such as the main construction sector, the plumbing sector or the glazier industry. In particular, our study might explain the descriptive evidence for the main construction sector which suggests that the MW regulations in that sector also led to a wage compression at the upper tail (Apel et al., 2012). Moreover, smaller firms in the plumbing and glazier industry, which face a similar market structure compared to the roofing sector, may react similarly to the national MW recently introduced by the German government and which is valid across all sectors. With 8.50€, the national MW has been set relatively high compared to international standards and might increase further.

Overall, the study demonstrates how institutions such as the MW that are mostly geared towards the lower rank may render unexpected side effects such as diminishing working conditions of workers higher up in the distribution that are mostly assumed to be unaffected by such policy reforms. Policy evaluators should thus take into account such heterogenous effects by looking at the entire distribution of earnings rather than simply looking at mean statistics. Furthermore, policy makers should be aware that a relatively high MW may more likely cause unfavourable effects in regions with lower price levels such as in East Germany as well as in smaller firms, especially in an economic downturn phase and may ultimately lead to decreasing returns to skills. However, further research is needed. First of all, the evaluation of employment spillovers that have been studied by Aretz et al. (2013) should be reconsidered by taking account of firm heterogeneities. It might also be interesting to follow the development in the sector in order to see how the distribution changes once the overall economy moves upward again, as the lastest figures in our data suggest. Finally, attempts to identify the cut-off point at which the MW induces unfavourable effects might be a fruitful line of further research.

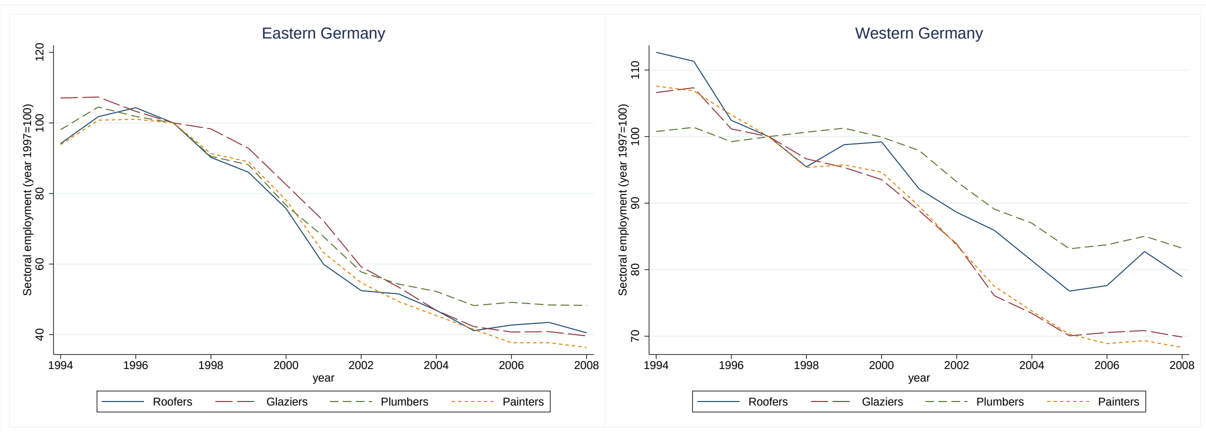
5.A Appendix

5.A.1 Development of gross daily wages and employment by sectors (BA data)

(a) Gross daily wages



(b) Employment



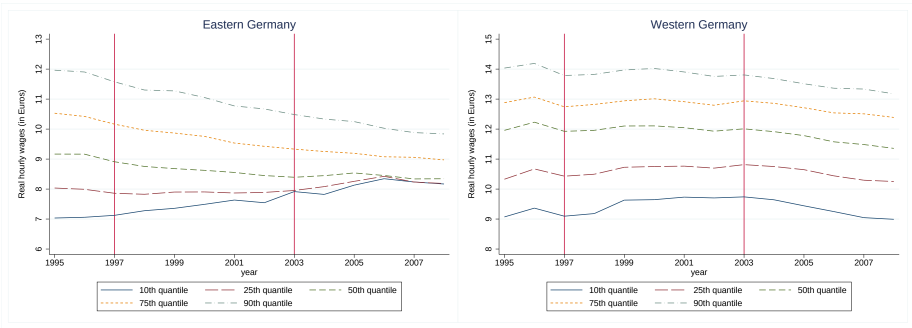
5.A.2 Detailed estimates for unconditional quantile regressions in Table 5.5 (Eastern Germany)

Dependent variable: log real daily wages	OLS	Quantile Regression Estimates					Inequality	
		$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$	Variance	Gini
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eastern Germany (N=451,245)								
Basic model:								
post-period interacted with treatment	3.74***	12.28***	9.44***	2.47***	-3.44***	-2.92***	-1.42***	-0.75***
	(27.34)	(62.93)	(75.72)	(20.49)	(-18.61)	(-9.89)	(-7.08)	(-35.66)
post-period dummy	-0.19***	-0.20***	-0.15***	-0.18***	-0.20***	-0.21***	-0.01**	-0.00
	(-86.48)	(-53.08)	(-70.46)	(-93.57)	(-77.73)	(-53.23)	(-2.73)	(-0.69)
treatment dummy	0.05***	0.05***	0.05***	0.06***	0.09***	0.03***	-0.01***	-0.00***
	(46.07)	(29.45)	(52.85)	(64.35)	(54.98)	(13.36)	(-5.73)	(-10.12)
age	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	-0.00*	-0.00***
	(70.17)	(52.39)	(62.11)	(66.25)	(61.26)	(43.41)	(-2.12)	(-5.58)
tenure in firm	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	-0.00***	-0.00***
	(116.00)	(83.16)	(96.76)	(111.82)	(84.01)	(41.39)	(-12.20)	(-29.38)
no high-school, with vocational traning degree	0.03***	0.05***	0.02***	0.01***	-0.00	-0.01*	-0.03***	-0.01***
	(11.59)	(14.21)	(9.83)	(5.75)	(-0.02)	(-2.11)	(-9.75)	(-21.27)
high-school, without vocational traning degree	-0.06**	-0.08***	0.00	0.02**	0.04**	0.06**	0.17***	0.02***
	(-2.80)	(-3.79)	(0.22)	(2.65)	(3.09)	(2.80)	(10.87)	(14.25)
high-school, with vocational traning degree	0.05***	0.07***	0.05***	0.05***	0.04***	0.04***	-0.02*	-0.00***
	(9.17)	(10.97)	(11.04)	(12.95)	(6.50)	(3.71)	(-2.37)	(-5.96)
technical university degree	0.02	0.03**	0.03***	0.02**	0.02*	-0.01	-0.00	-0.00
	(1.80)	(2.87)	(4.24)	(3.14)	(2.22)	(-0.30)	(-0.44)	(-1.74)
university degree	0.05*	0.00	0.02	0.01	0.06***	0.08*	-0.02	-0.00
	(2.50)	(0.08)	(1.29)	(1.30)	(3.34)	(2.16)	(-1.01)	(-0.15)
skilled worker	0.08***	0.09***	0.07***	0.07***	0.09***	0.08***	-0.01***	-0.00***
	(69.86)	(43.99)	(57.93)	(63.96)	(62.92)	(45.68)	(-6.87)	(-12.36)
master craftmen	0.33***	0.10***	0.13***	0.19***	0.39***	0.71***	0.14***	0.04***
	(114.90)	(35.96)	(75.91)	(117.74)	(134.87)	(106.67)	(42.43)	(112.71)
part-time dummy	-0.41***	-0.46***	-0.14***	-0.03***	0.03***	0.07***	0.58***	0.11***
	(-26.24)	(-32.30)	(-23.71)	(-7.98)	(4.59)	(7.63)	(65.81)	(114.16)
6-10 employees	0.03***	0.05***	0.03***	0.03***	0.02***	0.01***	-0.01***	-0.00***
	(35.68)	(31.16)	(34.10)	(29.77)	(18.51)	(8.24)	(-6.78)	(-17.56)
11-20 employees	0.05***	0.05***	0.04***	0.04***	0.04***	0.03***	-0.01***	-0.00***
	(47.18)	(30.91)	(41.91)	(41.92)	(32.41)	(19.21)	(-4.61)	(-11.61)
more than 20 employees	0.10***	0.08***	0.07***	0.08***	0.11***	0.14***	0.00**	0.00***
	(104.21)	(50.67)	(74.75)	(92.76)	(92.94)	(73.43)	(3.07)	(10.77)
number of low-skilled employees	-0.08***	-0.02***	-0.03***	-0.05***	-0.10***	-0.19***	-0.03***	-0.01***
	(-19.74)	(-3.90)	(-10.22)	(-15.91)	(-19.95)	(-22.29)	(-4.86)	(-15.54)
number of medium-skilled employees	-0.15***	-0.08***	-0.08***	-0.11***	-0.20***	-0.31***	-0.03***	-0.01***
	(-50.23)	(-20.82)	(-31.07)	(-45.27)	(-49.15)	(-44.13)	(-8.02)	(-25.31)
number of high-skilled employees	-0.01	-0.00	0.04***	0.04***	-0.05***	-0.13***	-0.03**	-0.01***
	(-1.41)	(-0.03)	(4.61)	(5.28)	(-4.91)	(-7.18)	(-2.94)	(-6.11)
constant	3.80***	3.44***	3.64***	3.82***	4.00***	4.27***	0.15***	0.05***
	(903.80)	(622.83)	(1084.92)	(1172.73)	(777.48)	(504.43)	(28.48)	(100.87)
N	451246	451246	451246	451246	451246	451246	451246	451246
R-squared	0.188	0.089	0.127	0.143	0.137	0.114	0.020	0.092
F	2618.4	954.8	2180.5	3224.1	2584.9	1142.5	292.3	1420.9

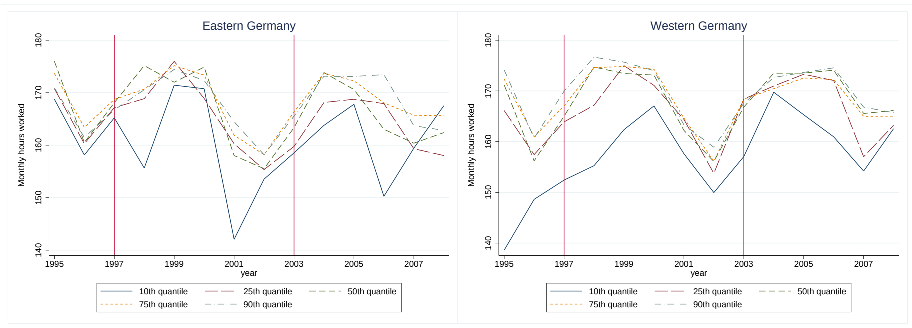
Notes: t-statistics in parenthesis. Significance levels: * 5%, ** 1%, *** 0.1%. Robust standard errors for OLS estimates. Standard errors of RIF regressions are bootstrapped with 50 replications. All models include a set of individual-level covariates including age, tenure, educational attainment (6 categories), occupational status (3 categories), part-time dummy as well as firm characteristics including qualification of workforce (3 categories) and company size (4 categories) and year dummies.

5.A.3 Development of real hourly wages, real monthly wages and hours worked (LAK data)

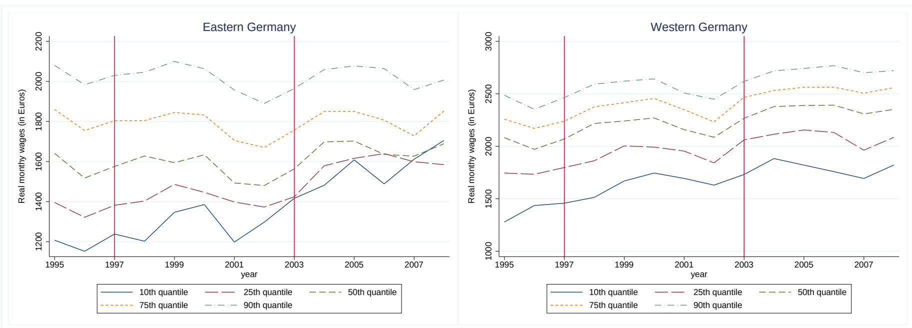
(a) Real hourly wages



(b) Monthly hours worked



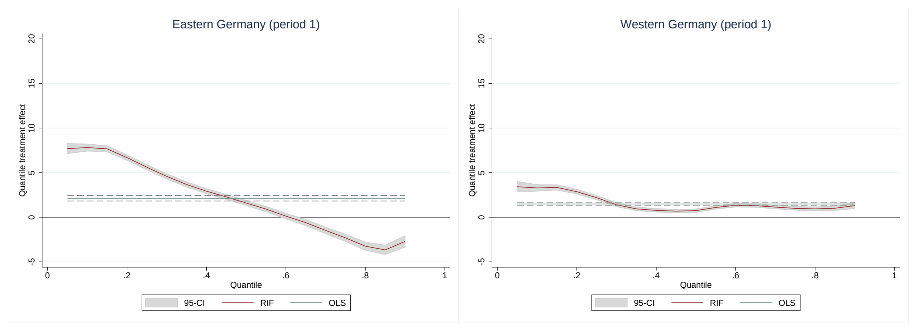
(c) Real monthly wages



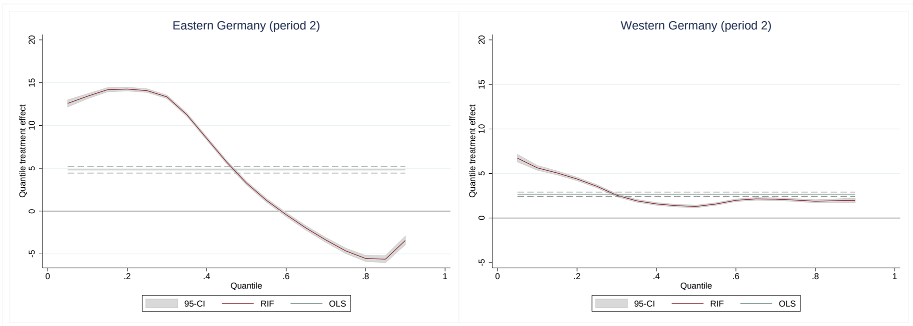
Notes: The vertical lines represent mark the introductions of the MW in October 1997 and the national MW in March 2003.

5.A.4 Unconditional Quantile regression estimates (interacted with period dummies)

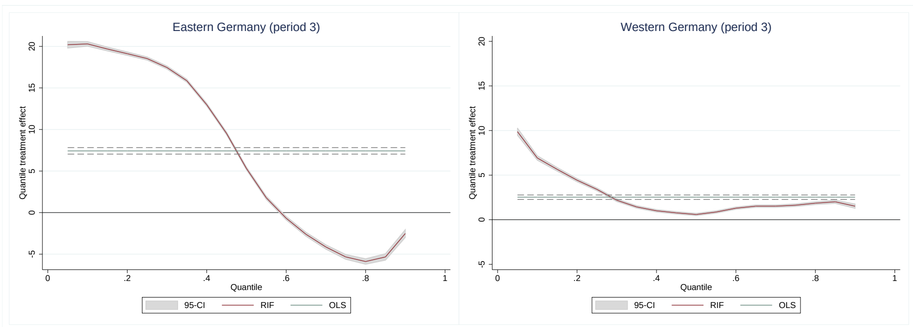
(a) 1998-2002



(a) 2003-2005



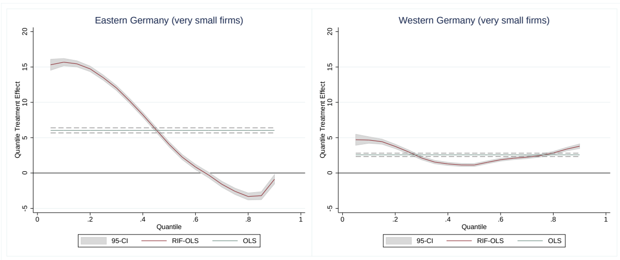
(a) 2006-2008



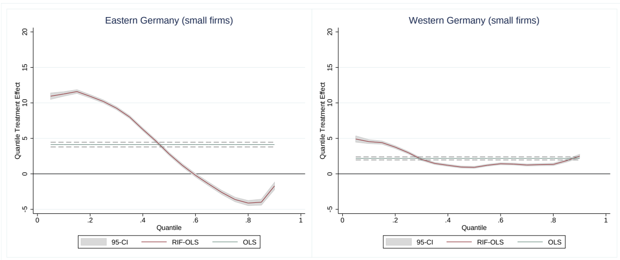
Notes: This figure shows OLS and QTE estimates including their 95% confidence intervals based on bootstrapping with 100 replications.

5.A.5 Unconditional Quantile regression estimates (interacted with firm size)

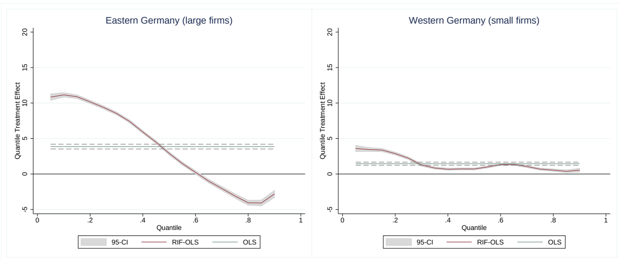
(a) very small firms



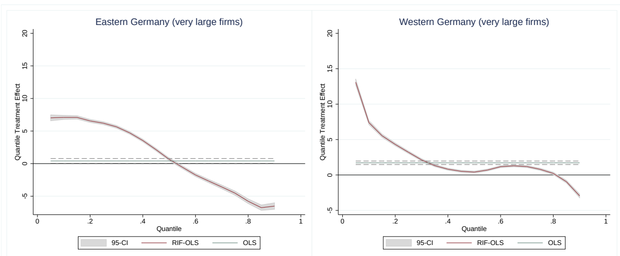
(a) small firms



(a) large firms



(a) very large firms



Notes: This figure shows OLS and QTE estimates including their 95% confidence intervals based on bootstrapping with 100 replications.

5.A.6 Conditional Quantile regression estimates

Dependent variable: log real daily wages	OLS	Quantile Regression Estimates				
		$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)
Western Germany (N=1,055,380)						
Basic model						
1998-2008	2.03***	2.64***	1.90***	2.02***	1.78***	0.74***
	(22.69)	(15.90)	(20.17)	(25.79)	(20.54)	(6.06)
Interacted with subperiods						
1998-2002	1.49***	1.88***	1.42***	1.44***	1.26***	0.55***
	(14.46)	(10.03)	(13.22)	(16.15)	(12.79)	(3.92)
2003-2005	2.68***	3.06***	2.26***	2.68***	2.45***	1.46***
	(22.16)	(13.47)	(17.32)	(24.77)	(20.45)	(8.62)
2006-2008	2.53***	3.78***	2.60***	2.66***	2.20***	0.33
	(19.80)	(16.28)	(19.48)	(24.01)	(17.93)	(1.92)
Interacted with firm size						
less then 6 employees	-2.19***	-2.55***	-1.40***	-0.18	-0.06	-1.70***
	(-19.13)	(-12.72)	(-12.16)	(-1.88)	(-0.56)	(-11.49)
6-10 employees	1.70***	2.74***	1.78***	1.67***	1.35***	-0.23
	(16.43)	(13.63)	(15.47)	(17.64)	(13.12)	(-1.58)
11-20 employees	3.66***	5.03***	3.19***	2.73***	2.47***	1.35***
	(34.72)	(24.08)	(26.67)	(27.75)	(23.02)	(8.75)
more than 20 employees	6.56***	6.79***	4.75***	4.52***	5.08***	5.36***
	(56.28)	(28.30)	(34.58)	(40.04)	(41.14)	(30.24)
Eastern Germany (N=451,245)						
Basic model						
1998-2008	3.74***	8.83***	5.58***	3.00***	0.62***	1.31***
	(27.34)	(47.02)	(42.01)	(24.50)	(3.76)	(5.32)
Interacted with subperiods						
1998-2002	2.12***	6.05***	2.94***	1.21***	-0.15	0.30
	(13.80)	(29.80)	(19.81)	(8.76)	(-0.84)	(1.07)
2003-2005	4.81***	11.20***	7.39***	3.75***	0.33	1.77***
	(25.92)	(40.47)	(36.54)	(19.96)	(1.33)	(4.66)
2006-2008	7.43***	15.20***	11.28***	6.79***	2.76***	3.44***
	(37.18)	(52.47)	(53.25)	(34.48)	(10.53)	(8.66)
Interacted with firm size						
less then 6 employees	0.61***	6.04***	3.89***	1.03***	-1.97***	-1.45***
	(3.64)	(26.40)	(23.48)	(6.67)	(-9.42)	(-4.66)
6-10 employees	2.62***	8.58***	5.11***	1.92***	-1.31***	-1.08***
	(17.00)	(38.15)	(31.34)	(12.61)	(-6.40)	(-3.53)
11-20 employees	3.60***	9.06***	5.58***	2.74***	-0.19	-0.01
	(22.69)	(39.05)	(33.21)	(17.46)	(-0.90)	(-0.02)
more than 20 employees	6.08***	10.09***	7.17***	5.01***	3.33***	4.25***
	(34.20)	(40.65)	(39.90)	(29.87)	(14.70)	(12.63)

Notes: t-statistics in parenthesis. Significance levels: * 5%, ** 1%, *** 0.1%. Robust standard errors for OLS estimates. Standard errors of RIF regressions are bootstrapped with 50 replications. All models include a set of individual-level covariates including age, tenure, educational attainment (6 categories), occupational status (3 categories), part-time dummy as well as firm characteristics including qualification of workforce (3 categories) and company size (4 categories) and year dummies.

Conclusion and Outlook

Part I. In contrast to many claims that location does not matter in a globalized and highly connected world with decreasing transport costs and improved communication technologies, this dissertation provides evidence that rather the opposite seems to be true, particularly in a knowledge-based economy. Regional differences in economic activity are largely due to an innovation sector that is gaining in importance and that tends to be more concentrated geographically compared to traditional industries. This is causing a polarization between regions. Whereas innovation hubs are increasingly able to attract innovative firms and high skill concentrations, other regions are left with economically depressed industries and a low-educated workforce. Overall, this thesis demonstrates that the well-being in a knowledge-based economy increasingly depends on where you live rather than your personal characteristics. More generally, inequality in advanced economies to a large extent reflects a geographical divide.

As demonstrated for the case of Germany, workforce ageing may reinforce this geographical divide. The reason for this is that demographic forces operating in the background of many advanced economies are largely geographically uneven. While urban areas are able to hold their average workforce age fairly constant by attracting young and skilled workers, due to advances related to thick labour markets and amenities, rural regions are suffering from depopulation and outmigration of their youngest workers. The findings suggest that it is very unlikely for such regions to reverse this trend due to strong contagion forces and cluster-wise path dependencies, and where it is not only a region's own history that influences its chances of modifying its status quo in the future, but where its surrounding environment plays an important role in limiting (favouring) the range of possible future outcomes. In particular, geographical areas facing a downward trend in terms of innovation performance and an ageing workforce are likely to pull other regions down as well. One reason for such geographical contagion forces might be age-selective migration that impacts the demographic structure of neighbouring regions through social networks and interregional ties. Clearly, such hypothesis could be tested more explicitly in future studies.

Regional planners aiming at reducing the demographically-driven geographical divide should take into account the role of agglomeration and contagion forces in their policy strategies. For instance, regions could cooperate more with other neighbouring regions in shaping an attractive metropolitan area for young workers, rather than competing against them. This might include alliances in the education system such as different universities with different specialisations (with respect to the field of study), but which are complementary. Similar joint initiatives may be conducted with respect to location policies, public (transport) infrastructure, cultural events and related fields. From a national perspective, financial grants may be attributed to metropolitan areas rather than single communities to increase incentives for such cooperation. Moreover, promoting innovation activities in beacon regions (for instance in Eastern Germany) with the aim of exploiting knowledge spillovers and (positive) agglomeration externalities might be more promising for the overall economy than turning around the trend in depopulated and less attractive rural areas.

In addition to the role of spatial dependencies, this thesis also stresses the importance of knowledge spillovers in fostering the local innovation sector. In general, social interactions among creative workers have shown to generate learning opportunities that enhance innovation and productivity. Such externalities thereby arise through face-to-face interaction within and between firms. The present dissertation shows that such learning opportunities are also relevant in the demographic context where younger and older workers benefit from knowledge exchange. The idea is that fluid abilities (speed of problem-solving and abstract reasoning) are known to decrease at older ages, whereas crystallized abilities (ability to use skills, knowledge and experience) remain at high functional levels until late in life. In fact, the thesis suggests that younger workers (under 30) and older workers (over 50) are complementary in the production of knowledge. However, the findings indicate an overall declining knowledge production in the future if demographic ageing further increases the size of the older workforce at the expense of the younger one. This is because it necessitates a sufficient size of the younger talent pool to benefit from the experienced and innovation-enhancing effect of older cohorts. Given the current demographic structure in Germany, the positive effects arising from complementarities between younger and older workers do not seem to suffice to compensate for the age-driven disadvantage in the generation of knowledge in case of a continued process of demographic ageing.

The findings stress not only the relevance of attracting a young talent pool from abroad. Regional policy makers should also be aware of the complementarities in the generation of ideas. In particular, they should foster knowledge exchange between different age groups in their local economy. Such strategies might include the support and promotion of knowledge cafés and social events where experienced stagers and young professionals network together to exchange ideas. The concept of business angels could be one example for an innovation-enhancing complementary between experienced business men and young entrepreneurs with promising ideas. Moreover, firm concepts that encourage age-diverse work teams and integrate the experience of older workers in the inventive process including part-time work concepts for the elderly could be promising in this regard. After all, exploiting the potentials of older workers might counterbalance potential negative effects from individual ageing.

In addition to the role of demographic ageing and regional migration in shaping regional disparities, further important trends might be investigated in future research. One of the trends that the author of the present dissertation has started to look into is technological change. Recent studies have highlighted the role of Task-Biased Technological Change (TBTC) in explaining the changing employment and wage structures of advanced economies (for a recent overview, see Autor 2013). Compared to former explanations that focused on relative changes in skill demand caused by technological change, known as the Skill-Biased Technological Change (SBTC) hypothesis, the TBTC literature focuses on changes in the demand for tasks. The central argument of this literature is that technological progress, in particular the diffusion of information and communication technology (ICT), increases labour demand for non-routine (cognitive and manual) tasks relative to labour demand for routine tasks, both cognitive and manual. In particular, routine tasks are those that can be automated because they follow a set of protocols which make them programable. Non-routine tasks, on the other hand, require the use of non-scriptable processes or person-to-person interaction. Cognitive non-routine tasks are mainly performed by high-skilled labour (e.g. managers, lawyers, etc.) and manual non-routine tasks by low-skilled labour (e.g. hairdressers, nurses, etc.). This leads to a polarization between worker groups because routine tasks are not uniformly distributed over the wage structure. In particular, jobs that are low- and high-paid are intense in non-routine tasks, while middling jobs are intense in routine tasks.

Several studies highlight the pervasive effect that technological change then has on these different occupations. However, most of these studies neglect the role of geography. Yet, just like routine tasks are concentrated in certain types of jobs, they are also spatially concentrated within countries. For instance, different regions may have different task intensities due to different sectoral specializations. Moreover, urban areas may constitute suitable environments for producing goods and services which require high shares of non-routine interactive tasks relative to rural areas. These different task intensities imply that regions will be differently affected by TBTC. The aim of future work in this field is therefore to show to what extent technological change can actually explain changes in regional labour market disparities in Germany and Europe (see Gregory et al. 2014).

Part II. In addition to the role of geography for labour market inequality, this thesis also deals with the effects of wage-setting institutions targeted at reducing inequality. In particular, the impact of the MW on both employment and earnings (inequality) are investigated in the German roofing sector, where the MW bites particularly hard. Whereas only 5% of the workforce earned a wage below the (next) MW level in Western Germany, almost half the workforce was affected in the eastern part of the country. Measured by the Kaitz Index, i.e. the ratio of the MW level and the median wage, that is ca. 1% in Eastern Germany, the bite has to be considered exceptionally high even by international standards (compare Figure 4 in the introduction).

The results demonstrate the importance of spillover effects of MW policies. In particular, the results suggest poorer chances of remaining employed in the roofing sector for all workers along the wage distribution. Although the deteriorating employment probabilities are strongest among the lowest deciles, upper-quantile workers also experienced lower prospects of continued employment. Similarly, the findings suggest large heterogeneities along the wage distribution regarding the MW effects on earnings. In particular, whereas lower-decile workers experienced large real daily wage increases that ripple up to workers above the median earner, the MW caused a reduction in real daily wages for upper quantile workers, where upper quantile workers are shown to comprise mostly skilled and experienced workers that have been working in the sector ever since.

One implication of this findings is that a MW may, under certain circumstances, affect the average pay reward for skilled labour. In the German roofing sector, the MW induced wage moderation among higher-skilled workers was the result of several occurrences. First of all, the increasing bite of the MW strongly raised the labour cost burden of firms, particularly for smaller ones. In fact, an accumulated labour cost share index suggests that the labour cost burden increased strongly during the observation period. Together with the overall downward trend in the construction sector, this seems to have limited the scope for firms to increase salaries of their skilled employees. Incentives for human capital investments have thus been shrinking in the sector and might also explain rising labour shortages that firms are facing recently as reported by sector insiders. Of course, it might be interesting to follow other sectors in Germany or countries with a similar setting in order to see how this finding can be generalized.

Overall, the thesis highlights the need for a broader perspective on employment and wage effects of MWs by also taking a closer look at workers who do not appear to be affected by the MW at a first glance. Moreover, the results put doubt on any attempts to identify employment effects of MWs by comparing workers with and without a binding MW within a covered sector. In the end, the size of the firm seems to matter very much, especially in a context like the German roofing sector where firms have relatively few employees and where the share of MW workers amounts up to 70%. Smaller firms have only a limited influence on market prices (price-takes) and have less opportunities for substituting labour with capital and may thus start compensating an increasing cost burden with wage moderation among their highest earning employees. Policy makers should thus be aware that a relatively high MW may more likely aggravate unfavourable effects in smaller firms rather than larger firms, especially in a phase of economic decline. Further research is definitely required. It is important that the evaluation of employment spillovers in this thesis be reconsidered and further analysed by taking firm heterogeneities into account. In addition, it is important to follow future developments and trends in the sector in order to determine how the distribution changes once the overall economy moves upward again. Finally, the impact of MWs and, more generally, wage-setting institutions on wage inequality is still not fully understood and because all indications are that they are a promising root cause, further research in the future is required (see also Autor et al. 2008, 2014).

Bibliography

Abowd, J. and Killingsworth, M. (1981). Structural models of the effects of minimum wages on employment by age groups. *Report of the minimum wage study commission*, 5:143–169.

Abowd, J. M., Kramarz, F., Margolis, D. N., and Philippon, T. (2000). The tail of two countries: Minimum wages and employment in France and the United States. IZA Discussion Papers 203.

Abraham, F. (1996). Regional adjustment and wage flexibility in the European Union. *Regional Science and Urban Economics*, 26(1):51–75.

Acs, Z., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, 31(7):1069–1085.

Acs, Z. and Audretsch, D. (1990). *Innovation and small firms.* MIT Press.

Acs, Z., Audretsch, D., and Feldman, M. (1992). Real effects of academic research: Comment. *American Economic Review*, 82(1):363–367.

Angrist, J. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis*, 27(2):93–115.

Anselin, L. (1996). The moran scatterplot as an ESDA tool to assess local instability in spatial association. *Spatial Analytical Perspectives on GIS*, 4:121.

Antonczyk, D., DeLeire, T., and Fitzenberger, B. (2010). Polarization and rising wage inequality: Comparing the US and Germany. ZEW Discussion Papers 10-015.

Apel, H., Bachmann, R., Bender, S., vom Berge, P., Fertig, M., Frings, H., König, M., Paloyo, A., Schaffner, S., Tamm, M., et al. (2012). Arbeitsmarktwirkungen der Mindestlohneinführung im Bauhauptgewerbe. *Journal for Labour Market Research*, 45(3-4):257–277.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–97.

Aretz, B., Arntz, M., Gottschalk, S., Gregory, T., Niefert, M., Rammer, C., Schröder, H., and Schütz, H. (2011). ZEW-Studie zu den Wirkungen eines Mindestlohns in der Dachdeckerwirtschaft. Research report for the Federal Ministry of Labour and Social Affairs.

Aretz, B., Arntz, M., and Gregory, T. (2012a). The minimum wage affects them all: Evidence on employment spillovers in the roofing sector. ZEW Discussion Paper 12-061.

Aretz, B., Arntz, M., and Gregory, T. (2013). The minimum wage affects them all: Evidence on employment spillovers in the roofing sector. *German Economic Review*, 14(3):282–315.

Aretz, B., Arntz, M., Gregory, T., and Rammer, C. (2012b). Der Mindestlohn im Dachdeckerhandwerk: Auswirkungen auf Beschäftigung, Arbeitnehmerschutz und Wettbewerb. *Journal for Labour Market Research*, 45(3-4):233–256.

Arntz, M. (2010). What attracts human capital? Understanding the skill composition of interregional job matches in Germany. *Regional Studies*, 44(4):423 – 441.

Arntz, M. and Gregory, T. (2014). What old stagers can teach us. Examining age complementarities in regional innovation systems. ZEW Discussion Paper 14-050.

Arntz, M., Gregory, T., and Lehmer, F. (2014). Can regional employment disparities explain the allocation of human capital across space? *Regional Studies*, 48(10):1719–1738.

Audretsch, D. and Feldman, M. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3):630–640.

Audretsch, D. and Feldman, M. (2004). Knowledge spillovers and the geography of innovation. *Handbook of Regional and Urban Economics*, 4:2713–2739.

Autor, D. (2013). The "task approach" to labor markets: An overview. *Journal for Labour Market Research*, 46:185–199.

Autor, D., Manning, A., and Smith, C. L. (2014). The contribution of the minimum wage to us wage inequality over three decades: A reassessment. Unpublished.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Backes-Gellner, U. and Veen, S. (2013). Positive effects of ageing and age diversity in innovative companies–large-scale empirical evidence on company productivity. *Human Resource Management Journal*, 23(3):279–295.

Bacolod, M., Blum, B. S., and Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, 65(2):136–153.

Baddeley, M. C., Martin, R. L., and Tyler, P. (2000). Regional wage rigidity: The European Union and United States compared. *Journal of Regional Science*, 40(1):115–141.

Berry, C. R. and Glaeser, E. L. (2005). The divergence of human capital levels across cities. *Papers in Regional Science*, 84(3):407–444.

Bertoli, S., Brücker, H., and Fernández-Huertas Moraga, J. (2013). The European crisis and migration to Germany: Expectations and the diversion of migration flows. IZA Discussion Papers 7170.

Blanchard, O. J. and Katz, L. F. (1992). Regional evolutions. *Brookings Papers on Economic Activity*, 23(1992-1):1–76.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115 – 143.

Bodvarsson, Ö. B. and Van den Berg, H. (2013). *The Economics of Immigration: Theory and Policy*, chapter Who Immigrates? Theory and Evidence, pages 79–106. Springer, New York.

Bönte, W., Falck, O., and Heblich, S. (2009). The impact of regional age structure on entrepreneurship. *Economic Geography*, 85(3):269–287.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review*, 77(4):531–53.

Borjas, G. J. (1991). Immigration and self-selection. In *Immigration, Trade and the Labor Market*, NBER Chapters, pages 29–76.

Borjas, G. J., Bronars, S. G., and Trejo, S. J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics*, 32(2):159–185.

Börsch-Supan, A. and Weiss, M. M. (2011). Productivity and age: Evidence from work teams at the assembly line. Mea discussion paper.

Bratsberg, B., Ragan Jr, J. F., and Warren, J. T. (2003). Negative returns to seniority: New evidence in academic markets. *Industrial and Labor Relations Review*, 56(2):306–323.

Brücker, H. and Trübswetter, P. (2007). Do the best go west? An analysis of the self-selection of employed East-West migrants in Germany. *Empirica*, 34(4):371–395.

Brown, C. (1999). Minimum wages, employment, and the distribution of income. *Handbook of Labor Economics*, 3:2101–2163.

Brunow, S. and Hirte, G. (2006). Age structure and regional economic growth. *Review of Regional Research*, 26:3–23.

Buch, T., Hamann, S., Niebuhr, A., and Rossen, A. (2014). What makes cities attractive? The determinants of urban labour migration in Germany. *Urban Studies*, 51(9):1960–1978.

Burda, M. and Wyplosz, C. (1992). Human capital, investment and migration in an integrated Europe. *European Economic Review*, 36(2-3):677–684.

Burda, M. C. and Hunt, J. (2001). From reunification to economic integration: Productivity and the labor market in Eastern Germany. *Brookings Papers on Economic Activity*, 32(2):1–92.

Caimcross, F. (1997). *The death of distance*. Harvard Business School Press.

Cameron, C., Gelbach, J., and Miller, D. (2011). Robust inference with multi-way clustering. *Journal of Business and Economic Statistics*, 29(2):238–249.

Card, D. and Krueger, A. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–793.

Card, D. and Krueger, A. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Reply. *American Economic Review*, 90(5):1397–1420.

Cardoso, A. R., Guimarães, P., and Varejão, J. (2011). Are older workers worthy of their pay? An empirical investigation of age-productivity and age-wage nexuses. *De Economist*, 159(2):95–111.

Ceh, B. (2001). Regional innovation potential in the United States: Evidence of spatial transformation. *Papers in Regional Science*, 80(3):297–316.

Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy*, 113(2):239–280.

Chiswick, B. R. (1999). Are immigrants favorably self-selected? IZA Discussion Paper 131.

Ciccone, A. and Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review*, 86(1):54–70.

Currie, J. and Fallick, B. (1996). The minimum wage and the employment of youth - Evidence from the NLSY. *Journal of Human Resources*, 31(2):404–428.

Dahl, G. B. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica*, 70(6):2367–2420.

Dall'erba, S. (2005). Distribution of regional income and regional funds in Europe 1989–1999: an Exploratory Spatial Data Analysis. *Annals of Regional Science*, 39(1):121–148.

Damm, A. P. and Rosholm, M. (2010). Employment effects of spatial dispersal of refugees. *Review of Economics of the Household*, 8(1):105–146.

Decressin, J. (1994). Internal migration in West Germany and implications for East-West salary convergence. *Review of World Economics*, 130(2):231–256.

Decressin, J. and Fatás, A. (1995). Regional labor market dynamics in Europe. *European Economic Review*, 39(9):1627 – 1655.

Dolton, P. and Bondibene, C. (2011). An evaluation of the international experience of minimum wages in an economic downturn. Research report for the low pay commission, Royal Holloway, University of London.

Dolton, P., Bondibene, C., and Wadsworth, J. (2011). Employment, inequality and the UK national minimum wage over the medium-term. *Oxford Bulletin of Economics and Statistics*, 74(1):78–106.

Dostie, B. (2011). Wages, productivity and aging. *De Economist*, 159(2):139–158.

Drews, N. (2008). Das Regionalfile der IAB-Beschäftigtenstichprobe. FDZ Datenreport 2, Institute for Employment Research, Nuremberg.

Dube, A., Lester, T., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *Review of Economics and Statistics*, 92(4):945–964.

Dube, A., Naidu, S., and Reich, M. (2007). The economic effects of a citywide minimum wage. *Industrial and Labor Relations Review*, 60(4):522–543.

Ederveen, S., Nahuis, R., and Parikh, A. (2007). Labour mobility and regional disparities: The role of female labour participation. *Journal of Population Economics*, 20(4):895–913.

Ertur, C. and Koch, W. (2006). Regional disparities in the European Union and the enlargement process: An Exploratory Spatial Data Analysis, 1995–2000. *Annals of Regional Science*, 40(4):723–765.

Etzo, I. (2011). The determinants of the recent interregional migration flows in Italy: A panel data analysis. *Journal of Regional Science*, 51(0):948–966.

European Parliament (2007). Regional disparities and cohesion - What strategies for the future. Report of the european parlament, Policy Department Structural and Cohesion Policies.

Eurostat (2013). Eurostat regional yearbook 2012. Publications office of the European Union.

Fairris, D. and Bujanda, L. (2008). The dissipation of minimum wage gains for workers through labor-labor substitution: Evidence from the Los Angeles living wage ordinance. *Southern Economic Journal*, 75(2):473–496.

Falk, A., Fehr, E., and Zehnder, C. (2006). Fairness perceptions and reservation wage - The behavioral effects of minimum wage laws. *Quarterly Journal of Economics*, 121(4):1347–1381.

Fazio, G. and Lavecchia, L. (2013). Social capital formation across space: Proximity and trust in European regions. *International Regional Science Review*, 36(3):296–321.

Feldman, M. (1999). The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Economics of Innovation and New Technology*, 8(1-2):5–25.

Feldman, M. P. and Florida, R. (1994). The geographic sources of innovation: Technological infrastructure and product innovation in the United States. *Annals of the Association of American Geographers*, 84(2):210–229.

Feyrer, J. D. (2008). Aggregate evidence on the link between age structure and productivity. *Population and Development Review*, 34:78–99.

Fields, G. S. (1976). Labor force migration, unemployment and job turnover. *Review of Economics and Statistics*, 58(4):407–415.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Fischer, M. and Varga, A. (2003). Spatial knowledge spillovers and university research: Evidence from Austria. *Annals of Regional Science*, 37(2):303–322.

Fitzenberger, B. (2012). Expertise zur Entwicklung der Lohnungleichheit in Deutschland. Working Paper 04, German Council of Economic Experts.

Fitzenberger, B. and Wilke, R. A. (2010). Unemployment durations in West Germany before and after the reform of the unemployment compensation system during the 1980s. *German Economic Review*, 11(3):336–366.

Florida, R. (2002). *The rise of the creative class and how its transforming work, leisure, community & everyday life*. Basic Books. New York.

Florida, R. and Gates, G. (2003). *The City as an Entertainment Machine: Research in Urban Policy*, volume 9, chapter Technology and tolerance: The importance of diversity to high-technology growth, pages 199–219. Elsevier.

Florida, R., Mellander, C., Stolarick, K., and Ross, A. (2012). Cities, skills and wages. *Journal of Economic Geography*, 12(2):355–377.

Fratesi, U. and Riggi, M. R. (2007). Does migration reduce regional disparities? The role of skill-selective flows. *Review of Urban and Regional Development Studies*, 19(1):78–102.

Friedman, T. (2005). *The World is Flat: A Brief History of the Globalized World in the Twenty-first Century*. Penguin UK.

Frings, H. (2013). The employment effect of industry-specific, collectively bargained minimum wages. *German Economic Review*, 14(3):258–281.

Fritsch, M. and Slavtchev, V. (2007). Universities and innovation in space. *Industry & Innovation*, 14(2):201–218.

Frosch, K. and Tivig, T. (2009). Age, human capital and the geography of innovation. In Kuhn, M. and Ochsen, C., editors, *Labour Markets and Demographic Change*, pages 137–146. VS Verlag fuer Sozialwissenschaften.

Gartner, H. (2005). The imputation of wages above the contribution limit with the German IAB employment sample. FDZ Methodenreport 2, Institute for Employment Research, Nuremberg.

Gernandt, J. and Pfeiffer, F. (2009). Wage convergence and inequality after unification: (East) Germany in transition. *Labor Markets and Economic Development*, 73:387–404.

Giese, E. (2002). The spatial pattern of invention activities in West Germany 1992-1994. In *Technological change and regional development in Europe*, pages 9–27. Physica Verlag, Heidelberg, New York.

Giese, E. and von Reinhard Stoutz (1998). Indikatorfunktion von Patentanmeldungen für regionalanalytische Zwecke in der Bundesrepublik Deutschland. *Spatial Research and Spatial Planning*, 56(5-6):414–420.

Giesecke, J. (2009). Socio-economic risks of atypical employment relationships: Evidence from the German labour market. *European Sociological Review*, 25(6):629–646.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–52.

Göbel, C. and Zwick, T. (2012). Age and productivity: Sector differences. *De Economist*, 160(1):35–57.

Gregory, T. and Patuelli, R. (2013). Regional age structure, human capital and innovation - Is demographic ageing increasing regional disparities? ZEW Discussion Paper 13-057.

Gregory, T., Salomons, A., and Zierahn, U. (2014). Technological change and regional labor market disparities in Europe. Unpublished.

Griliches, Z. (1979). Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics*, 10(1):92–116.

Griliches, Z. (1987). *R&D, patents and productivity.* University of Chicago Press.

Griliches, Z. (1998). Patent statistics as economic indicators: A survey. In *R&D and Productivity: The Econometric Evidence*, pages 287–343. University of Chicago Press.

Grossman, J. (1983). The impact of the minimum wage on other wages. *Journal of Human Resources*, 18(3):359–378.

Haltiwanger, J. C., Lane, J. I., and Spletzer, J. R. (1999). Productivity differences across employers: The roles of employer size, age, and human capital. *American Economic Review*, 89(2):94–98.

Havnes, T. and Mogstad, M. (2014). Is universal child care leveling the playing field? *Journal of Public Economics*, forthcoming.

Hellerstein, J. K., Neumark, D., and Troske, K. R. (1996). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. NBER Working Papers 5626.

Helpman, E., Itskhoki, O., and Redding, S. (2010). Unequal effects of trade on workers with different abilities. *Journal of the European Economic Association*, 8(2-3):421–433.

Hethey-Maier, T. and Seth, S. (2010). The establishment history panel (BHP) 1975-2008: Handbook version 1.0. 2. FDZ Datenreport 4, Institute for Employment Research, Nuremberg.

Hierro, M., Maza, A., and Villaverde, J. (2013). A proposal for detecting spatial contagion: Some evidence on the international migration distribution in Spain. *Papers in Regional Science*, 92(4):811–829.

Horn, J. L. and Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26:107–129.

Hsing, Y. (2000). On the substitution effect of the minimum wage increase: New evidence. *Applied Economics Letters*, 7(4):225–228.

Hunt, G. L. and Mueller, R. E. (2004). North American migration: Returns to skill, border effects, and mobility costs. *Review of Economics and Statistics*, 86(4):988–1007.

Hunt, J. (2004). Are migrants more skilled than non-migrants? Repeat, return, and same-employer migrants. *Canadian Journal of Economics*, 37(4):830–849.

Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association*, 4(5):1014–1037.

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review*, 79(5):957–970.

Jones, B. F. (2010). Age and great invention. *Review of Economics and Statistics*, 92(1):1–14.

Juhn, C., Murphy, K. M., Topel, R. H., Yellen, J. L., and Baily, M. N. (1991). Why has the natural rate of unemployment increased over time? *Brookings Papers on Economic Activity*, 22(2):75–142.

Kanbur, R. and Rapoport, H. (2005). Migration selectivity and the evolution of spatial inequality. *Journal of Economic Geography*, 5(1):43–57.

Katz, L. F. and David, H. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, 3:1463–1555.

Kline, P. (2010). Place based policies, heterogeneity, and agglomeration. *American Economic Review*, 100(2):383–387.

König, M. and Möller, J. (2009). Impacts of minimum wages: A micro data analysis for the German construction sector. *International Journal of Manpower*, 30(7):716–741.

Koenker, R. (2005). *Quantile regression.* Number 38. Cambridge University Press.

Koenker, R. W. and Bassett, G. W. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Kosfeld, R. and Werner, A. (2012). Deutsche Arbeitsmarktregionen - Neuabgrenzung nach den Kreisgebietsreformen 2007-2011. *Spatial Research and Spatial Planning*, 70(1):49–64.

Kraft, K., Rammer, C., and Gottschalk, S. (2012). Minimum wages and competition: The case of the german roofing sector. ZEW Discussion Papers 12-083, Center for European Economic Research.

Kulu, H. (2008). Fertility and spatial mobility in the life course: Evidence from Austria. *Environment and Planning A*, 40(3):632–652.

Lallemand, T. and Rycx, F. (2009). Are older workers harmful for firm productivity? *De Economist*, 157(3):273–292.

Le Gallo, J. (2004). Space-time analysis of GDP disparities among European regions: A Markov chains approach. *International Regional Science Review*, 27(2):138–163.

Le Gallo, J. and Ertur, C. (2003). Exploratory Spatial Data Analysis of the distribution of regional per capita GDP in Europe. *Papers in Regional Science*, 82:175–201.

Lee, D. S. (1999). Wage inequality in the United States during the 1980s: Rising dispersion or falling minimum wage? *Quarterly Journal of Economics*, 114(3):977–1023.

Lehman, H. C. (1953). *Age and achievement.* Princeton University Press.

LeSage, J. and Pace, R. (2009). *Introduction to spatial econometrics.* Chapman & Hall/CRC, Boca Raton.

Leuvensteijn, M. V. and Parikh, A. (2002). How different are the determinants of population versus labour migration in Germany? *Applied Economics Letters*, 9(11):699–703.

Lindh, T. and Malmberg, B. (1999). Age structure effects and growth in the OECD, 1950-1990. *Journal of Population Economics*, 12(3):431–449.

Linneman, P. (1982). The economic impacts of minimum wage laws: A new look at an old question. *Journal of Political Economy*, 90(3):443–69.

Lucas, R. J. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Machin, S., Manning, A., and Rahman, L. (2003). Where the minimum wage bites hard: Introduction of minimum wages to a low wage sector. *Journal of the European Economic Association*, 1(1):154–180.

Machin, S. and Van Reenen, J. (2008). Changes in wage inequality. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

Malecki, E. (1997). *Technology and economic development: The dynamics of local, regional and national competitiveness*. Addison Wesley Longman.

Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Manski, C. (2000). Economic analysis of social interations. *Journal of Economic Perspectives*, 14(3):115–136.

Mauro, P. and Spilimbergo, A. (1999). How do the skilled and the unskilled respond to regional shocks? the case of Spain. *IMF Staff Papers*, 46(1):1–17.

McCormick, B. (1997). Regional unemployment and labour mobility in the UK. *European Economic Review*, 41(3-5):581–589.

Mertens, A. (2002). Regional and industrial wage dynamics in West Germany and the United States. *Journal of Economics and Statistics*, 222(5):584–608.

Möller, J. and Tubadji, A. (2009). The creative class, bohemians and local labor market performance: A micro-data panel study for Germany 1975–2004. *Journal of Economics and Statistics*, 229(2-3):270–291.

Moretti, E. (2011). Local labor markets. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, Part B, chapter 14, pages 1237 – 1313. Elsevier.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.

Morrison, P. S. (2005). Unemployment and urban labour markets. *Urban Studies*, 42(12):2261–2288.

Nahuis, R. and Parikh, A. (2002). Factor mobility and regional disparities; East, West, home's best? CPB Discussion Paper 4.

Neumark, D., Schweitzer, M., and Wascher, W. (2004). Minimum wage effects throughout the wage distribution. *Journal of Human Resources*, 39(2):425–450.

Neumark, D. and Wascher, W. (1994). Employment effects of minimum and subminimum wages: Reply to Card, Katz, and Krueger. *Industrial and Labor Relations Review*, 47(3):497–512.

Neumark, D. and Wascher, W. (1995). Minimum-wage effects on school and work transitions of teenagers. *American Economic Review*, 85(2):244–249.

Neumark, D. and Wascher, W. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Comment. *American Economic Review*, 90(5):1362–1396.

Neumark, D. and Wascher, W. (2008). *Minimum wages*. MIT Press.

Nickell, S. and Bell, B. (1995). The collapse in demand for the unskilled and unemployment across the OECD. *Oxford Review of Economic Policy*, 11(1):40–62.

Niebuhr, A., Granato, N., Haas, A., and Hamann, S. (2012). Does labour mobility reduce disparities between regional labour markets in Germany? *Regional Studies*, 46(7):841–858.

O'brien, R. (1992). *Global financial integration: The end of geography.* Council on Foreign Relations Press New York.

Ochel, W. (2005). Decentralizing wage bargaining in Germany - A way to increase employment? *Labour*, 19(1):91–121.

OECD (2005). How persistent are regional disparities in employment? The role of geographic mobility. In *OECD Employment Outlook 2005*, chapter 2.

Oster, S. M. and Hamermesh, D. S. (1998). Aging and productivity among economists. *Review of Economics and Statistics*, 80(1):154–156.

Parikh, A. and Leuvensteijn, M. V. (2003). Interregional labour mobility, inequality and wage convergence. *Applied Economics*, 35(8):931–941.

Patacchini, E. and Rice, P. (2007). Geography and economic performance: Exploratory Spatial Data Analysis for Great Britain. *Regional Studies*, 41(4):489–508.

Patuelli, R., Griffith, D. A., Tiefelsdorf, M., and Nijkamp, P. (2012). Spatial filtering methods for tracing space-time developments in an open regional system: Experiments with German unemployment data. In A. Frenkel, P. N. and McCann, P., editors, *Societies in Motion: Regional Development, Industrial Innovation and Spatial Mobility*, pages 247–68. Edward Elgar Publishing.

Peri, G. (2005). Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87(2):308–322.

Pettengill, J. (1981). The long run impact of a minimum wage on employment and the wage structure. *Report of the Minimum Wage Study Commission*, 6:63–104.

Pissarides, C. A. and McMaster, I. (1990). Regional migration, wages and unemployment: Empirical evidence and implications for policy. *Oxford Economic Papers*, 42(4):812–31.

Poot, J. (2008). Demographic change and regional competitiveness: The effects of immigration and ageing. *International Journal of Foresight and Innovation Policy*, 4(1):129–145.

Prskawetz, A., Mahlberg, B., and Skirbekk, V. (2007). Firm productivity, workforce age and educational structure in Austrian industries in 2001. In Clark, R., Hiro, O., and Mason, A., editors, *Population Aging, Intergenerational Transfers and the Macroeconomy*. Edward Elgar Publishing.

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography*, 2(4):373–406.

Puhani, P. A. (2001). Labour mobility: An adjustment mechanism in Euroland? Empirical evidence for Western Germany, France and Italy. *German Economic Review*, 2(2):127–140.

Rattenhuber, P. (2014). Building the minimum wage: The distributional impact of Germanys first sectoral minimum wage on wages and hours across different wage bargaining regimes. *Empirical Economics*, 46(4):1429–1446.

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, 34(3):380 – 400.

Rey, S. (2001). Spatial empirics for economic growth and convergence. *Geographical Analysis*, 33(3):195–214.

Rey, S. J., Murray, A. T., and Anselin, L. (2011). Visualizing regional income distribution dynamics. *Letters in Spatial and Resource Sciences*, 4(1):81–90.

Riddell, W. C. and Song, X. (2011). The impact of education on unemployment incidence and re-employment success: Evidence from the U.S. labour market. *Labour Economics*, 18(4):453 – 463.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–37.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Sato, R. and Koizumi, T. (1973). On the elasticities of substitution and complementarity. *Oxford Economic Papers*, 25(1):44–56.

Scherer, F. (1983). The propensity to patent. *International Journal of Industrial Organization*, 1(1):107–128.

Schneider, L. (2008). Aging and technological innovativeness - A linked employer-employee analysis. *Comparative Population Studies*, 33(1):37–54.

Schöb, R. and Wildasin, D. E. (2007). Economic integration and labor market institutions: Worker mobility, earnings risk, and contract structure. *Regional Science and Urban Economics*, 37(2):141–164.

Südekum, J. (2008). Convergence of the skill composition across german regions. *Regional Science and Urban Economics*, 38(2):148–159.

Senftleben, C. and Wielandt, H. (2012). The polarization of employment in German local labor markets. Humboldt University zu Berlin Discussion Paper 2012-013.

Simonton, D. K. (1988). Age and outstanding achievement: What do we know after a century of research? *Psychological Bulletin*, 104(2):251.

Skirbekk, V. (2004). Age and individual productivity: A literature survey. *Vienna Yearbook of Population Research*, 1(1):133–153.

Todaro, M. P. (1969). A model for labor migration and urban unemployment in less developed countries. *American Economic Review*, 59(1):138–48.

Van Ours, J. C. and Stoeldraijer, L. (2011). Age, wage and productivity in Dutch manufacturing. *De Economist*, 159(2):113–137.

Von Hippel, E. (1994). 'Sticky information' and the locus of problem solving: Implications for innovation. *Management Science*, 40(4):429–439.

Wilke, R. A. (2005). New estimates of the duration and risk of unemployment for West-Germany. *Journal of Applied Social Science Studies*, 125(2):207–237.

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data.* MIT Press.

This cumulative dissertation reflects on labour market inequalities from two perspectives. The first deals with the role of geography in shaping inequalities in an aging knowledge-based economy. In particular, it explores recent trends in demographic aging for regional labour markets in Germany, investigates the causal link between an aging workforce and regional innovation and examines the role of selective migration in shaping regional disparities. The results demonstrate that location matters in an aging economy where regions are increasingly becoming polarized due to agglomeration forces and urbanisation trends. However, the thesis also shows opportunities for regions to enhance their innovation performance by exploiting knowledge externalities between young and older workers. The second perspective deals with the economic effects of minimum wages as one important policy instrument targeted at reducing inequalities. In particular, it investigates both employment effects as well as the effects on earnings and wage inequality of a sectoral minimum wage in Germany, where the minimum wage bites extraordinary hard by international standards. The results show that a minimum wage geared towards low-wage workers may not only increase their own earnings, but also render unexpected side effects for workers located higher up in the wage distribution, including reduced employment chances among qualified workers and diminished returns to skills.