ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Ehrl, Philipp

Working Paper

A breakdown of residual wage inequality in Germany

BGPE Discussion Paper, No. 150

Provided in Cooperation with:
Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Bavarian Graduate Program in Economics (BGPE)

Suggested Citation: Ehrl, Philipp (2014) : A breakdown of residual wage inequality in Germany, BGPE Discussion Paper, No. 150, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bavarian Graduate Program in Economics (BGPE), Nürnberg

This Version is available at:
https://hdl.handle.net/10419/99980

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

BGPE
Bavarian Graduate Program in Economics

BGPE Discussion Paper

No. 150

A breakdown of residual wage inequality in Germany

Philipp Ehrl

May 2014

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.
Friedrich-Alexander-University Erlangen-Nuremberg
© Philipp Ehrl

B G P E Discussion Paper Series

A breakdown of residual wage inequality in Germany

Philipp Ehrl*
University of Würzburg‡

May 2014

Abstract

The present paper applies several regression-based decomposition methods to analyze the impact of region-, worker-, firm- and sector-specific determinants on the wage level and the continuous increase in wage inequality between 1995 and 2007 in Germany. In contrast to prior studies, more than 50% of the wage dispersion and almost the entire increase in wage inequality are explained in this approach. Altogether, the entire growth of wage dispersion occurs within regions and changes in the composition of wage determinants are minor compared to changes in their returns. I find that occupational attributes are the most important wage determinant. Changes in the firm size premium in combination with assortative matching also depress wages in the bottom of the distribution while they increase wages at the top. Workers with an unemployment record or an occupation in the service, construction and logistics sectors particularly experience falling wages.

Keywords: wage inequality, skill biased technical change, inequality decompositions
JEL Classification: J31, J24, R12

*I would like to thank Michael Pflüger, Klaus Gründler, Fabian Kindermann, Josef Seibold, Georg Struch and seminar participants at the University of Würzburg for fruitful discussions and comments. Furthermore, I thank the Research Data Center (RDC) of the German Federal Employment Agency at the Institute for Employment Research for the data access, especially Simon Trenkle for his support. The data basis of this publication is the SIAB. Data access was via guest research spells at the FDZ and via controlled data remote access.

‡Faculty of Economics, Sanderring 2, 97070 Würzburg, Germany. *philipp.ehrl@uni-wuerzburg.de*

1 Introduction

A marked increase in wage inequality has been observed in Germany, the US and other industrialized countries over the past 30 years, regardless of whether wage inequality is measured as the dispersion of wages, the skill premium, etc. (Acemoglu and Autor 2011; Berman *et al.* 1998; Dustmann *et al.* 2009; Juhn *et al.* 1993; Lindley and Machin 2013; Van Reenen 2011). Three reasons for changes at the bottom of the wage distribution are identified: positive supply shocks of low-skilled workers, a decrease of the minimum wage and declines in unionization. However, most of the rising inequality could not be explained, i.e., had to be attributed to the residual wage. The residual wage is defined as the wage after controlling for the observable variables in the study, which typically do not go beyond a subset of education, age, tenure, gender and industry-specific variables. All of the above cited papers conclude that the large and increasing amount of residual wage inequality is due to skill-biased technical change (SBTC). Lemieux (2006) criticizes this assessment because residual wage growth may either be caused by returns to unobservable skills, their composition in the labor force or measurement error. Above all, Lemieux (2006) finds support that US wage inequality increased mainly due to changes in composition. Autor *et al.* (2008) relativize his result by noting that aggregate inequality measures disguised a polarization of the wage distribution. These gains in the upper and lower part of the wage distribution relative to the middle are consistent with a "richer version" of SBTC. The advancing information technology requires and complements abstract tasks, which are found especially in high-wage occupations, whereas it substitutes routine tasks that are more present in occupations in the middle of the wage distribution.

The present paper provides an overview and a distinction of the scope of several wage determinants that are considered separately in other recent studies on inequality. In contrast to prior studies, I additionally relate worker's unobservable adherent skills to occupational attributes which are useful in examining prior conclusions about SBTC more precisely. Consequently, my definition of residual wages is much more narrow than in earlier studies, so that explainable wage determinants instead of residual wages are in the center of attention here. In addition, most of those papers concern US wages, so in a sense, I extend the well-known study by Dustmann *et al.* (2009) to explain the rise in Germany's wage inequality between 1993 and 2010 using more wage determinants, more recent data that covers East and West Germany.[1] Aside from individual characteristics, I include diverging returns in cities of different size (Baum-Snow and Pavan 2013), the regional price level (Moretti 2013) and the matching between employer and employees (Card *et al.* 2013).[2]

[1] Wage inequality is also an interesting object of study because Biewen and Juhasz (2012) show that labor income is the main driver of the overall income inequality, leaving only a smaller role to changes in household characteristics or in the tax and transfer system.

[2] While the Stolper-Samuelson framework and its rigid distinction between high- and low-skilled workers have received little support concerning the labor market trends over the past 20 years, more recent tests of heterogeneous firm models and offshoring have. Helpman *et al.* (2014) structurally estimate a version of Melitz (2003) with search frictions that provides a link between the residual wage between firms and firm size, exporter status and assortative matching. Another trade related channel that recently received great attention is import and export exposure. Autor *et al.* (2013) demonstrate that final goods imports from China have adverse wage effects on workers in competing domestic industries. The analysis in Dauth *et al.*

Building on Juhn *et al.* (1993) and Morduch and Sicular (2002), I apply regression-based decompositions that allow *side-by-side comparisons* of the impact of the various wage components. Both papers focus on percentile ratios and the variance, whereas I extend their approach to the entire distribution of wages. The idea is very simple. I estimate a reduced form wage regression that encompasses the wage determinants mentioned above. Due to the additive linearity of the regression, aggregate inequality indices, as well as the entire wage distribution, can be decomposed exactly into these different wage components. The share that each component adds to the wage inequality measure is interpreted as its impact. The approach is flexible enough to incorporate two more sub-decompositions: a distinction of the wage dispersion between and within regions, and a distinction between changes in the composition of wage determinants and changes in their returns. Yet, decomposition methods generally fail to include any general equilibrium effects.

Figure 1 presents the core facts about the development of the German (inflation-adjusted) wage distribution from two different points of view. Panel (a) shows the change in log wage along its distribution, separately for the periods 1995–2001 and 2001–2007. Inequality increases in both periods but the divergence of labor incomes within the period of 12 years is impressive. Individuals at the bottom of the distribution lose up to 15%, while workers at the top gain up to 20%. Panel (b) shows the development of the 80th, 50th and 20th percentile of the wage distribution *within* regions. The divergence of the mean of the regions' 80th and 20th percentiles reflect the increase in inequality. However, the spread of the confidence intervals around these means is more or less constant. Both graphs already preview some important results and indicate that the development of the German labor market is essentially different from the one in the US. First, there is no wage polarization in Germany because workers with the lowest wages do not gain with respect to workers in the middle of the distribution. Second, the dispersion of wages *between* regions is quite constant over time. Consequentially, the rise in inequality comes entirely from changes within regions.

[Insert figure 1 about here.]

Another key result is that the wage determinants considered in this paper explain about 75% of the total wage change. This is remarkable given that most preceding studies conclude that the distribution of residual wages essentially resembles that of raw wages. These observable factors include education, age, tenure, plant size, dummies for regions of different size, the interaction of these agglomeration indicators with the former characteristics, a regional price index, industrial affiliation and dummies for 341 different occupations. In contrast to Lemieux (2006), I find that changes in the composition of these characteristics play a minor role compared to changes in their returns. In the following, I discuss these wage determinants separately and compare my findings to the previous literature.

The most important contribution to the development of wages comes from occupations. Returns in technical and business jobs grow in particular, even though their supply in-

(2014) suggests that once export opportunities are considered as well, these adverse effects may be offset. Unfortunately, data constraints do not allow the inclusion trade related variables in a satisfactory way.

creases as well. These occupations involve a large amount of analytical and interactive skills (Autor *et al.* 2003; Spitz-Oener 2006), which confirms that the rise in wage inequality is indeed consistent with SBTC. The difference is, that this conclusion is no longer based on unobserved factors. Notwithstanding, a nuanced view of SBTC – as articulated in Autor *et al.* (2008) – is required because highly educated workers do not benefit across-the-board.

The fact that Germany does not experience a polarization of wages similar to the US rests on several factors that depress wages in the bottom of the distribution. I identify that these workers are on average younger, work in smaller firms and are more likely to have an unemployment record. The returns to these characteristics fall over time. Furthermore, low-wage workers in the service and construction sector continuously lose in real terms.[3] Wage declines in the lower third of the German wage distribution over the 1990s are already well documented in Dustmann *et al.* (2009). The sources identified in their paper, namely, the increased supply of low-skilled workers and the decline in unionization, complement my findings.[4] However, they apply a different econometric approach where only one variable of interest is changed to simulate counterfactual wage distributions. Moreover, only education and age are controlled for, so that again the main part of the changes in the upper part of the wage distribution is embodied in the residual wage.

The link between wage inequality and remuneration differences across employers and regions have received much less attention. In my regression-based approach, I identify wage premiums for observably equivalent workers in larger firms, in large cities, diverging returns to observable characteristics across regions, and wage compensations due to cost of living disparities. All of these spatial disparities account for only 13% of the wage dispersion between regions and also add little explanation of wage changes. My approach reveals that the effective wage compensations for higher costs of living are only 20%. Note the conceptual difference to Moretti (2013) who deflates wages with a regional price index and finds that the resulting differences in well-being are less pronounced than nominal wage differences.[5] Furthermore, even though productive workers and firms are sorted into large regions and even though the share of skilled workers and their wage premium grows over time, the *growth* of inequality is evenly spread across regions, as anticipated in figure 1. That is, the degree of regional sorting does not intensify. In contrast, the employment share and premium for high-skilled workers increase substantially between regions in the US.[6] The associated expansion of low-skilled (service) jobs in these regions also contributes

[3] Another possible reason for the polarization of wages and employment is offshoring. Like SBTC, the relocation of production particularly substitutes workers with routine tasks in manufacturing industries, whereas the tasks of managers are more required (Acemoglu and Autor 2011). Empirical evidence for the associated wage effects is found in the US by Ebenstein *et al.* (2014) as well as in Germany by Baumgarten *et al.* (2013).

[4] Glitz (2012) documents that after the fall of the Iron Curtain almost 3 million ethnic Germans migrated to Germany.

[5] In a related study, Blien *et al.* (2009) analyze the German rural-urban wage gap using a regression approach. However, they rely on imputed regional price indices, for which they report a much larger explanatory power.

[6] Although Baum-Snow and Pavan (2013) observe that inequality growth is correlated with the size of the city, sorting of workers with high observable and unobservable skills into large cities does not attenuate, in line with my findings. Lindley and Machin (2014) stress that the inequality growth between regions is demand-induced and that it is not particularly concentrated in large cities but in regions where

to the observed labor market polarization (Lindley and Machin 2014).

Regarding the matching pattern, I find that high-wage workers are employed in large firms. At the same time, these plants pay a wage premium to all workers. The fact that this firm-specific remuneration differential is growing significantly over time explains about 6% of the change in the total wage gap and up to 7% of the total wage dispersion. Again, it proves insightful to separate between the change in the allocation of workers to firms and the change in the wage premium. It becomes apparent that the firm-specific part of the increase in inequality is almost entirely due to changes in the return to firm size. It fits the picture that the pattern of assortative matching has no particular relation to the size of the region (Ehrl 2014), whereas a positive relation is found in the US by Andersson *et al.* (2007). Lehmer and Möller (2010) also stress the importance of the firm size premium in explaining the rural-urban wage gap in Germany. Their approach involves following cohorts of workers, but still unobservable differences account for more than one half of the wage gap. Helpman *et al.* (2014) separate the industry and occupation specific part of the wage dispersion from a firm-specific part. The former is about 30% and thus comparable to my results, while the latter is about 3 times larger than what I capture by firm size. They then show that a heterogeneous firm model that includes exporters and assortative matching provides a good fit to the firm-specific part of the wages in Brazilian data. This study is static and does not consider the change in wage inequality. It is also not comparable to my paper because the model as a whole is evaluated, not its separate components. Card *et al.* (2013) analyze the German matching pattern but focus on the impact of unobservables captured by worker and firm fixed effects. The dispersion captured by both dummies rises and their assortative matching with each other also contributes to wage inequality.

The remainder of the paper is composed of three parts. Section 2 describes the data. Section 3 discusses various different decomposition methods, their application and results. I begin with the estimation of the basic wage regression, then discuss simple aggregate inequality decompositions and finally analyze changes along the wage distribution. Section 4 concludes the paper.

2 Data

This study uses the SIAB (Sample of Integrated Labour Market Biographies) provided by the Research Data Center of the German Federal Employment Agency. The SIAB is a 2% random sample of all employees subject to social security in Germany. It is an administrative dataset which is based on mandatory annual reports to the social security agencies, thus the sample is representative, large and highly reliable. Another advantage is that it provides comprehensive information about the workers as well as some characteristics of the employer, such as its size, the sector and the location of the plant. For a detailed description of this data set, see vom Berge *et al.* (2013). A prior version of this data is used in the well-known study on wage inequality by Dustmann *et al.* (2009).

computerization and R&D is higher. These are regions that were high-skill abundant before.

I constrain the sample on full-time employment relations of men at a specific day in each year between 1993 and 2010. Beginning in 1993 allows the inclusion of workers in East and West Germany. I focus on June 30th because the firm-level information is valid only for this day. This also avoids problems associated with wage calculations if a person does not have the same job throughout the year. Unfortunately, top-coding of wages is a common problem with the German administrative data. Such top-coded wages are imputed using the predictions of a censored regression model, similar to Dustmann *et al.* (2009) and Card *et al.* (2013). These imputed daily wages are then deflated by the national CPI. For details about the imputation procedure, sample selection, the description of further variables and summary statistics, see the data Appendix.

Additional information about local prices and employment densities from the Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) is used. The BBSR distinguishes nine different county types according to the county's population density, size and centrality. I aggregate these county types into six categories to control for different wage levels and returns in counties of different size and density. The classification also considers whether counties are close to agglomerations, to account for the spatial diffusion of agglomeration economies beyond administrative borders. An overview of these county types and their distribution is provided in the following figure.

[Insert figure 2 about here.]

The BBSR surveyed a total of 7.3 million single commodity and housing prices at the level of counties between 2006 and 2008 to compute comprehensive and representative local price indices. Its computation uses the weighting scheme and basket as the national CPI, cf. Kawka *et al.* (2009) for the details. Housing costs account for 20% of the price index and comprise rents as well as the rental value of owned housing. The inclusion of housing costs is important for the accuracy of the approach and both measures are preferable over housing values (Winters 2009). This is the first and only survey of regional prices that covers all counties in Germany. Because the BBSR provides only one value of the price index per county, which is valid for the years 2006 to 2008, an implicit assumption for its application in the following panel regression is that price level differences are constant throughout the entire period of this study. However, the results are not sensitive to the inclusion of these local price indices.

3 Decomposition methods and their application

The aim of the empirical analysis is to evaluate the impact of separate wage determinants. In a first step, a wage regression provides estimates for the effects of the components. A component effect is defined by the value of the variable $\mathcal{X}$ multiplied by its coefficient β, e.g., $\beta\mathcal{X}$. The same way that all component effects add up to the wage in the linear regression equation, I will decompose the value of wage inequality measures into shares that add up exactly and reflect the importance of each component. The remainder of

this section describes different decomposition methods and their empirical results. A short overview of the applied decomposition methods is given in subsection 3.2. Subsection 3.3 uses the variance to measure inequality, 3.5 consider the Gini index, the Theil index and percentile ratios. Subsection 3.4 additionally examines the share of inequality between and within regions and how much the different wage determinants contribute to these two dimensions of inequality. Afterward, the decomposition is performed separately for each percentile of the wage distribution. Subsection 3.7 further separates the component effects into changes in returns and changes in the composition of workers' characteristics. I begin by describing the rationale for the inclusion of each of the wage determinants.

3.1 The basic wage regression

All decompositions in the present paper build on the separability of the wage in additive components, that are identified from the wage regression described by

$$\begin{aligned} w_{it} = \beta_0 &+ \beta_x X_{it} + \beta_z Z_{jt} + \beta_w West_{kt} + West_{kt} \cdot [\beta_{xw} X_{it} + \beta_{zw} Z_{jt}] + \beta_s D_s \\ &+ \beta_k D_k + \beta_p Pi_k + D_k \cdot [\beta_{xk} X_{it} + \beta_{zk} Z_{jt}] + \beta_t D_t + \upsilon_i D_i + \epsilon_{it} \end{aligned} \tag{1}$$

where w_{it} is the deflated imputed daily log wage of individual i in year t. Additionally, a worker i is characterized by three more dimensions: sector s and county k, where her employer j is located. Each capital letter in eq. (1) aggregates important determinants of the individual wage rate which are described and justified in the following. Most obvious and common are the personal characteristics included in X_{it}. Education, tenure and age determine productivity and are thus the benchmark for the workers's wage level. I distinguish between low, medium and high educational achievements and include the square of both tenure and age because productivity may eventually decrease before retirement.

The remuneration of workers across firms may differ for several reasons, some of which are linked to the firm's productivity and production function, and others stem from organizational, institutional or working conditions (Brown and Medoff 1989). In the absence of detailed plant-level data, size is a reasonable indicator for productivity (Idson and Oi 1999). The shares of low- and high-skilled workers provide an approximation to the firm's production function (Haltiwanger *et al.* 1999). Both shares, the employer's size and its square are denoted by Z_{jt}. Differences in wage levels and in the compensation of characteristics between East and West Germany, are captured by a West indicator variable and its interaction with X_{it} and Z_{jt}. Additional dummies for having experienced a period of unemployment or having completed an apprenticeship are also included in eq. (1).[7] Those workers are likely to have a lower productivity and accept lower wage offers to (re-)enter the labor market (Mortensen 1987).

Sizable inter-industry wage differentials exist but can only be partly explained by different working conditions or the sorting of skilled workers (Krueger and Summers 1988). Overall

[7] These variables are only included in X_{it} but not in the interaction terms of X_{it} due to a low variation of these indicators.

differences between the 59 sectors and 18 years are captured by the dummies in D_s and D_t, respectively. Because Gibbons and Katz (1992) note that the size of the firm or the sectoral classification alone are incapable of explaining large wage differentials, it is informative to compare the predicting power of all these and other characteristics side-by-side.

In the spatial dimension, wages differ due to agglomeration externalities, amenities and living costs. From a theoretical point of view, only nominal wages may differ because in a spatial equilibrium without migration barriers utility levels are equal, cf. Roback (1982). Amenities such as a beautiful surrounding area or a pleasant climate attract people but populated regions are themselves more attractive because resident firms and workers are more productive, and thus better remunerated.[8] The balance is restored by the cost of living, principally through the housing market, see Kosfeld and Eckey (2010) and Suedekum (2006) for empirical and theoretical evidence. Elevated living costs raise local service prices and wages even for workers who do not directly benefit from agglomeration advantages. The local price and housing cost index (Pi_k) captures how workers are compensated for the cost of living. Note that I do not deflate wages directly because this paper evaluates wage determinants instead of utility differences as in Moretti (2013).[9] Remaining differences in amenities, agglomeration benefits (that accrue to all workers) and any other constant wage disparities in regions of different size are captured by county type dummies D_k, as defined in section 2 and shown in figure 2.[10] Agglomeration economies that benefit a specific skill group or groups of firms are captured via the interaction terms $D_k\left[\beta_{xk}X_{it} + \beta_{zk}Z_{jt}\right]$.

Finally, the term $\upsilon_i D_i$ captures all time-invariant worker-specific wage differences after controlling for all other observable characteristics in the regression. These unobservable differences are commonly attributed to adherent skills. There is no reason to believe that unobserved skills are uncorrelated with the individuals' education or with the quality of her employer and region by means of sorting. Consequently, those variables' coefficients and explanatory power are likely to be biased when unobservable skills are not controlled for by D_i. On the other hand, these adherent differences become part of the wage residual, on which conclusions about the SBTC have largely been based. In contrast, I exploit the panel dimension and include worker fixed effects D_i. At the same time, these dummies act as a placeholder in eq. (1), and thus other time-invariant determinants of the wage may be extracted from them. To do so, the estimated $\hat{\upsilon}_i$ are regressed on occupation and German citizenship dummies. Occupations represent adherent abilities that distinguish workers from each other. An engineer, for example, acquires good analytical and technical skills through his academic studies and professional experience, and such skills are highly

[8] A multitude of different agglomeration externalities are suggested in the literature, see e.g. the survey in Puga (2010). Ehrl (2013) and Baldwin *et al.* (2010) find evidence that the positive productivity effects are mainly transmitted via the labor market.

[9] Dividing wages by the local price index is equivalent to constraining the coefficient of Pi_k to 1 and implies that workers are fully monetarily compensated for higher living costs. It turns out that this is far from reality.

[10] Even with county fixed effects, the total explanatory power of the wage regression improves only slightly. Estimation with fixed effects for all counties is problematic in the presence of worker fixed effects because the former are identified only from the within variation of individuals. Constrained by computational power and the number of movers between counties, the interaction terms $D_k X_{it}$ and $D_k Z_{jt}$ would need to be abandoned.

demanded in a number of industries. Therefore, in this approach, occupations serve to visualize part of the unobserved skills of workers. The share of $\hat{\upsilon}_i$ in the total wage dispersion is typically about 60%, cf. e.g. Combes *et al.* (2008). At least the occupational choice explains about 36% thereof. This explained part is denoted by υ_{expl} and the residual in this second stage regression is υ_{res}. Note that these occupation dummies only capture average differences in labor market valued skills between occupations, but this is exactly the intention. Other time-variant and personal abilities acquired through on the job training, etc., are part of the wage residual that is defined in this study as $\xi = \upsilon_{res} + \epsilon_{it}$.

The estimation of eq. (1) is performed separately for three different periods: 1993–1998, 1999–2004, 2005–2010. It is an arbitrary compromise because there is tradeoff between the length and the number of periods. One the one hand, the longer the periods, the better the identification of all coefficients, especially the fixed effects. On the other hand, the more different estimations/periods, the more different coefficients are obtained and the better can the changes in returns over time be traced.

It is obvious that all of the dummy variables in the wage regression (D_s, D_k, D_t, D_i) need to have an omitted reference category, otherwise the column vectors of their matrices are linearly dependent. The categorical variables also enter the estimation as a set of dummy variables and thus require a reference category. This necessity sometimes complicates the interpretation of the inequality decompositions, as we shall see in the following.[11]

The bottom line is that not only discrimination but several economic fundamentals give rise to wage inequality. In the first place, wages differ because there are diverse types of skills, i.e., different labor input. Workers, firms and places possess varied productivities that create remuneration differentials. Such fundamental factors can be and have to be accounted for, before it is appropriate to start thinking about how inequalities can be reduced. Likewise, these factors may provide insights why the distribution of labor income has changed.

3.2 Overview of decomposition methods

This subsection briefly discusses the benefits and drawbacks of the decomposition methods used in the following to facilitate the overview of the type of information we can derive from the several approaches. Details of these methods can be found in the respective subsections. A cornerstone in the literature on the decomposition of inequality measures into different components is Shorrocks (1982). His paper discusses basic principles that any non-arbitrary inequality decompositions should satisfy. For example, the inequality measure should be symmetric, continuous and should yield a value of zero if income is equally distributed across the population. The decomposition should also be continuous, symmetric, independent of the aggregation level of the components, the sum of all contributions of income should exactly add up to the total inequality value, and the representation of the decomposition should be unique. In Shorrock's illustration, total income is separated into

[11] The reference worker in my estimation has medium education and works in the car repair industry in a county of the lowest density.

actual flows of wage income, capital income, etc. Morduch and Sicular (2002) transfer this concept to estimated flows from determinants of a single source of income, for example, the wage. The values of these flows correspond to the component effects identified in the prior wage equation eq. (1). When the components are correlated, the interaction effects of these components pose a potential problem to the decomposition. Shorrocks (1982) argues that the "natural" decomposition attributes one half of the interaction term to both of the correlated components. The variance is one of the few inequality measures for which such a natural decomposition is applicable that concurrently fulfills the desired properties.

In the same way, the variance of wages can first be separated into differences between and within regions. Then, both parts may be decomposed into the detailed wage determinants. The benefit of this exercise is to evaluate more precisely where and how the change in inequality occurs. Regional inequality continues to be an important topic in German and European politics, and the literature on this kind of regional wage inequality is scarce.

Yet the variance has three disadvantages. It is criticized because it does not satisfy the transfer axiom (Foster and Ok 1999).[12] Second, the variance is not scale invariant. Thus, it makes a difference if one uses daily, hourly or monthly wages. Third, comparability to other studies is limited because the variance is not a frequently used measure, at least in the literature on inequality. However, not every index is exactly decomposable into its factor components, as is the case for the variance, due to the possible interaction between components. Shorrocks (2013) provides a different approach called *Shapley value decomposition* which is applicable to any inequality measure. Its implementation is a little more complicated because components have to be considered sequentially and thus their contribution depends of the order in which they are considered (path dependency). Hence, all possible sequences have to be calculated to report the average contribution of each component. To keep the approach manageable, only a few components can be evaluated limiting its accuracy compared to the variance decomposition. The indices I decompose with the Shapley value approach are the Gini index, the Theil index, the 80/20 percentile ratio and the standard deviation of logs. A comparison between different inequality indices is generally advisable because due to their definition and theoretical grounding, the components' contributions differ across the indices.

With the indices considered so far, the inequality is expressed as a single value. This is clear and simple but it does not yield conclusions about how different groups in society are affected. That is why the entire wage decomposition is considered in the remaining part of the paper. First, I consider the change of each wage component separately in each percentile of the distribution. The analysis is refined by distinguishing between changes in returns ($\Delta\beta$) and changes in the composition of characteristics ($\Delta\mathcal{X}$), cf. eq. (3) below. Thereby supply and demand effects may be distinguished and the causes of wage inequality can be better understood.

To separate the coefficient from the composition effects, an additional counterfactual wage

[12] The transfer axiom states that a transfer from a richer to a poorer individual will unambiguously decrease inequality.

w_{aux} is required to perform the following decomposition of wage changes

$$w_{t=2007} - w_{t=1995} = (w_{t=2007} - w_{aux}) + (w_{aux} - w_{t=1995}) \tag{2}$$

$$= \underbrace{(\beta_{t=2007} - \beta_{t=1995})}_{\equiv\Delta\beta} \mathcal{X}_{t=2007} + \beta_{t=1995} \underbrace{(\mathcal{X}_{t=2007} - \mathcal{X}_{t=1995})}_{\equiv\Delta\mathcal{X}} \tag{3}$$

where the several coefficients and wage determinants from eq. (1) are represented in stacked form $\beta\mathcal{X}$ and the subscript t refers to an arbitrary year. In this case, it denotes a year in the middle of the first and the third period for which this regression is estimated. The counterfactual wage corresponds to a situation in which individuals were remunerated for their characteristics as in the year 1995, but the composition of their characteristics remains as it was in 2007. Therefore, the first term in both equations reflects the influence of changed returns and the second term is equal to the part of the wage change that is attributable to shifts in the composition of workers' characteristics. Eq. (3) follows because of the underlying linear wage regression, but the insight of this decomposition is general and applicable to other type of estimations.

Blinder (1973) applies this decomposition to the difference in the average wage between males and females. To extend this idea to the entire distribution, I follow Juhn *et al.* (1993) and construct the counterfactual wage $w_{aux} \equiv \beta_{t=1995}\mathcal{X}_{t=2007}$ by multiplying the average value of each variable in $\mathcal{X}_{t=2007}$ *in every percentile* with the respective coefficient estimated for the year 1995.[13] According to eq. (3), the impact of every wage determinant can be depicted along the wage distribution. DiNardo *et al.* (1996) develop a semi-parametric version of this decomposition. First, they estimate the density of wages conditional on some wage determinants. To construct the counterfactual w_{aux}, the conditional distribution of one of the wage determinants – union membership M_i in their application – is replaced by its distribution in another year (or group). This reweighting of the wage distribution requires an estimate of the distribution of M_i in both periods. The advantage of their counterfactual simulation is that, as in the literature on Propensity Score Matching, under certain assumptions, the change in M_i may be interpreted as a causal treatment effect.[14]

A combination of their method with the present fixed effects regression with numerous interaction terms is, however, beyond the scope of this paper. For the same reason it is not feasible to estimate eq. (3) using quantile regressions. Whereas quantile regressions provide separate coefficients, that make the illustration of the changes more accurate, I use coefficients that identify an average effect. The choice between these techniques and the fixed effects model is a tradeoff between the accurateness in different aspects. My approach yields unbiased coefficients because unobserved skills are controlled for and the additionally identified occupational attributes provide other valuable insights that cannot be found in other studies. Finally, it should be emphasized that the following assump-

[13] Juhn *et al.* (1993) take the wage distribution in the average of all years as their reference point and they exchange values of $\mathcal{X}$ that correspond to individuals with the same rank in the distribution. They then focus on a specific quantile ratio instead of the entire distribution. In fact, Juhn *et al.* (1993) require that the rank of each individual is unchanged in order to substitute the value of $\mathcal{X}$ with the one from the average distribution. Because I only focus only on percentiles, this need not necessarily be the case.

[14] In the case of DiNardo *et al.* (1996) the main interest is in the effect of the minimum wage.

tions are presupposed just like in the other available decompositions. General equilibrium effects are turned off, e.g., the coefficients of the counterfactual wage function w_{aux} are unaffected when the distribution of covariates is changed.[15] Besides, detailed regression based decompositions require the additive linearity of the wage function.

3.3 Variance decomposition

A first approximation of the importance of all components in the wage regression in eq. (1) is achieved through a variance decomposition. To simplify the notation, I define the following aggregate wage component estimates

$$w_{it} = level + X + Z + S + K_{\text{total}} + \upsilon_{\text{expl}} + \xi \tag{4}$$

$$\begin{aligned} \text{where:} \quad K_{\text{total}} &= K + Pi + XK + ZK \\ &\equiv \hat{\beta}_k D_k + \hat{\beta}_p Pi_k + D_k \cdot \left[\hat{\beta}_{xk} X_{it} + \hat{\beta}_{zk} Z_{jt}\right] \\ level &\equiv \hat{\beta}_0 + \hat{\beta}_t D_t \\ X &\equiv \hat{\beta}_x X_{it} + \hat{\beta}_w West_{kt} + \hat{\beta}_{xw} West_{kt} \cdot X_{it} \\ Z &\equiv \hat{\beta}_z Z_{jt} + \hat{\beta}_{zw} West_{kt} \cdot Z_{jt} \\ S &\equiv \hat{\beta}_s D_s \end{aligned} \tag{5}$$

The wage residual ξ was defined as $\upsilon_{res}+\epsilon_{it}$. For convenience, I denote the seven component estimates on the RHS of eq. (4) with W^c for the following derivation and suppress the lower case index, so that $\sum W^c$ is equal to the original wage w. The "natural" decomposition rule derived by Shorrocks (1982) for the variance of wages $\sigma^2(w)$ is thus given by

$$\begin{aligned} \sigma^2(w) &= \sum_{c=1}^{7} cov(W^c, w) \\ &= \sum_{c} \sigma^2(W^c) + \sum_{c \neq d} \sum_{d} corr(W^c, W^d)\sigma(W^c)\sigma(W^d) \end{aligned} \tag{6}$$

since the covariance can be expressed as

$$cov(W^c, w) = \sigma^2(W^c) + \sum_{c \neq d} corr(W^c, W^d)\sigma(W^c)\sigma(W^d) \tag{7}$$

That is, the interaction terms in eq. (7) are assigned equally to each of the components. Dividing eq. (6) by $\sigma^2(w)$ shows that the covariance between each component and the original wage reflects the relative importance of the components. The magnitude of $cov(W^c, w)$ has another intuitive interpretation. It is equal to the mean of "(A) the inequality which would be observed if income component c was the only source of income differences; and (B) the amount by which inequality would fall if differences in factor c income receipts were eliminated" (Shorrocks 1982: 209). This is not true for other decomposition methods

[15] Other extensions of the Oaxaca-Blinder decomposition and a detailed discussion of their required assumptions, advantages and limitations is given in Fortin *et al.* (2011).

of the variance and most other inequality indices (an obvious exception is the coefficient of variation $\sigma^2(w)/\mu$). Gibbons *et al.* (2013) argue that the choice for whichever assignment of the interaction terms is still arbitrary. Yet eq. (6) satisfies the axioms in Shorrocks (1982), and it is straightforward in the presence of more than two components.

Columns 1 and 2 in table 2 illustrate the importance of the wage components in absolute and relative terms, according to the variance decomposition in eq. (6). The values in panels 1 to 3 refer to the mean in each of the three periods. In the first line of each panel, the wage dispersion indicates that there is a steady increase in inequality in Germany. However, contrary to most of the studies, the share of the residual wage *decreases* from 54.9% to 46.2% indicating that the observable wage determinants in this paper account for more than half of the inequality level. The largest fraction is explained by occupational attributes. They dwarf the remaining wage components and even increase from 28% to 32%. The remaining worker- and firm-specific characteristics are the next largest portions. The absolute contribution to inequality of X doubled and of Z increased even more.

In this multivariate setting, the share of the sectoral component is below 4% and almost constant over time. Because the total wage dispersion grows, a constant share still indicates that the contribution of sectoral differences add to the rise in wage inequality. The estimated coefficient for the log of the local price index in the wage regression is approximately 0.2.[16] Hence, via their labor income, workers are only compensated for 20% of the local price level differences. This suggests that amenities compensate workers for the largest part of higher living costs. All regional components, including the price index, are less important than the sectoral differences. Combined, their explanatory power decreases to even less than zero. While the contributions of XK, K and the local prices to inequality are constant, albeit positive, the regional firm-specific differences ZK fall substantially and actually reduce inequality. This merely means that the remuneration in large regions *relative* to rural regions (the reference category) declines. Consequentially, the reasons for this finding may be either increased sorting of high-wage individuals to rural regions, higher firm-specific premiums in rural regions or lower premiums in agglomerations. An explicit distinction of disparities between and within regions in the next subsection sheds more light on this finding.

Fact 1. The wage determinants in this study explain more than 50% of the wage dispersion and about 75% of its change from 1995–2007. Occupational attributes have by far the largest explanatory power, followed by other personal characteristics and firm-specific wage premiums.

[Insert table 2 about here.]

[16] The regression output is omitted due to space constraints. Furthermore, most of the coefficients are not readily interpretable because of the various interaction terms. Therefore it is more informative to present the aggregated component effects.

3.4 Inequality between and within regions

This subsection estimates how much of the total inequality is attributable to regional differences and how such differences themselves are composed. It is important to note that regional differences arise for two reasons. They are either due to regional differences in the remuneration for personal and firm-specific characteristics XK and ZK, respectively, or because the composition of workers differs between regions. For example, if some regions have a larger share of high-skilled workers and there is a general skill premium, X contributes to the regional wage gap. On the other hand, K and Pi only capture regional wage differences across-the-board and thus they do not have variation within regions. In accordance with the observation level of the regional price indices and the classification of county type dummies D_k, I analyze the inequality within and between German counties.

First, I describe some important differences in the composition and the remuneration of the workforce across the six different county types used in this study, cf. figure 2 above. The columns in the following table 3 show county means in the first and the last period and their growth rate in this time interval. Columns 4 and 6 display the ratio of county means to the aggregate mean to readily quantify the amount of the disparities. As expected, average wages are highest in and around agglomerations (county types no. 1 and 2), as well as in core cities (no. 4), and are lowest in rural regions (no. 6). The difference between log wages in agglomerations and rural regions amounts to 0.26 and is constant over time. Thus without consideration of productive characteristics, the urban-rural wage gap is 26%. On the one hand, the regional disparities in the share of university graduates and in firm size are substantial as well. On the other hand, education premiums differ by less than 2% in period 1, which suggests the disparities in average wages are due to the sorting of workers and firms. The share of high-skilled workers grows in all of the county types and intensifies existing disparities. Nevertheless, the high-skill premium still differs by no more than 2–3%. This raw skill premium may either be due to differences in demand or due to other observable or unobservable worker attributes. In contrast, regional disparities in firm size attenuate over time because firms in agglomeration shrink by 31% and firms in rural regions grow by 10–13%. An exception are firms in core cities which are already on average the biggest and continue to expand.

[Insert table 3 about here.]

I extend the previous variance decomposition in eq. (6) and sub-divide each component effect in its dispersion within and between regions.

$$cov(W^c, w) = cov(W^c - \bar{W}^c, w) + cov(\bar{W}^c, w) \tag{8}$$

The first term reflects the variation of a wage component effect W^c within counties by subtracting its county mean $\bar{w}^c$. The inequality between counties is equal to the covariance between the county means and the original wages. Again, the decomposition into these two terms is exact.

Columns 3–6 in table 2 display the corresponding results. Overall inequality between counties in Germany (measured in terms of the variance of log wages) is constant at about 0.03. Its relative share, however, is decreasing from 20% to 14%, due to the growth in overall inequality. Hence, the increase inequality is entirely coming from the dispersion of wages within counties. Regarding the single wage determinants, the table reveals that their relative importance differs in both dimensions.

Between counties, the spatial distribution of occupation- and firm-specific payments, υ_{expl} and Z, increased inequality in absolute and relative terms. Yet the rise is offset by payments for individual characteristics and by the regional components K, Pi, XK, ZK. The negative impact of ZK on inequality in the final period suggests that firms impose lump-sum discounts in dense regions (where wages are highest) *relative* to workers in the reference category. This accords with the shrinking firm size in agglomerations displayed in table 3. Again, the low explanatory power of Pi indicates that workers are only compensated monetarily for higher living costs to a small degree. The largest fraction ($\approx 1/3$) of regional wage disparities is due to the distribution of the unobservable skills υ_{res}. Even though the sorting of workers with the highest education category into large regions intensifies, as shown in table 3, the education premium rises more or less evenly and the inequality due to all personal characteristics in X_{it} decreases slightly.

Within counties, the regional firm-specific payments also reduce inequality, as previously discussed. Quite the contrary, the nationwide employer-specific pay component adds to the increase in inequality. It will be revealed below whether this is due to low-wage workers being paid less or because high-wage workers receive more. Finally, the returns on education, age, tenure and occupation are identified as the main drivers of wage inequality in Germany.

The presented extension of the variance decomposition to accommodate an additional decomposition into subgroups (counties) is straightforward. Subgroup decompositions also exist for other inequality indices, such as the entropy measures by Theil. However, to date no combination of a decomposition by subgroups and by components is available for the Theil index. Nonetheless, to provide a more common inequality measure than the variance, I calculate the wage inequality within and between counties using the Theil-T index for two benchmark cases: Original wages and residual wages ($\tilde{w}_{it}$), i.e., the wage in the hypothetical situation where all workers have equivalent observable characteristics. I use the same formula as in Shorrocks and Wan (2005: 63),

$$\begin{aligned} T(w) &= \frac{1}{N}\sum_i \frac{w_i}{\mu} log\left(\frac{w_i}{\mu}\right) \\ &= \sum_k \left\{ \frac{N_k}{N}\frac{\mu_k}{\mu}\left[\frac{1}{N_k}\sum_{i\in k}\frac{w_i}{\mu_k} log\left(\frac{w_i}{\mu_k}\right)\right]\right\} + \sum_k \left\{\frac{N_k}{N}\frac{\mu_k}{\mu} log\left(\frac{\mu_k}{\mu}\right)\right\} \end{aligned} \tag{9}$$

where N is the number of all individuals and μ is defined as their mean wage. Likewise, N_k and μ_k apply to the individuals in county k. The first equation states the definition of the Theil-T index. The second equation shows how the index is decomposed into the

inequality within and between regions. It is easy to see that the first term in eq. (9) is a weighted sum of single Theil indices in each region. Analog to the variance decomposition in eq. (6), the second term is also weighted and only refers to the relation between the regional and the aggregate mean wages.

The compact representation of only two benchmark cases allows the development of the inequality for every year in the sample to be presented separately. Table 4 confirms most of the prior assessments. Wage inequality grows steadily and the increase stems purely from changes within regions. In 1995, the wage inequality between counties accounted for about 19% of total inequality, compared to 20% in terms of the variance. Extracting all explainable wage determinants and re-calculating the Theil index with the residual wages ($\tilde{w}_{it}$) reduces the level and change over time by more than 50%. This confirms that the unexplainable inequality share falls. Like in table 2, the residual wage differences between counties are almost constant and close to zero.

[Insert table 4 about here.]

The finding that the dispersion of productive abilities between regions does not contribute much to the overall level of inequality is in line with finding in other countries. Gibbons *et al.* (2013) find that even with regional fixed effects the share of inequality between regions amounts to a maximum of 6% in Britain. The survey by Shorrocks and Wan (2005) indicates that this share is somewhat higher in most developing and industrializing countries.

Evidence on regional wage inequality in Germany is scarce. Recent studies focus on the urban-rural wage gap. Blien *et al.* (2009) find that in data from 1993 almost 10 percentage points of the 25% urban wage premium is explained by higher housing and consumer prices in agglomerations. The remainder of the wage gap is explained by the dispersion of individual characteristics. Note that the wage gap in table 3 shows a similar wage premium of 26% and that this premium is on average constant until 2007. The difference is that Blien *et al.* (2009) use imputed price levels and divide the raw wage directly by the price index, whereas I identify the impact of all explanatory variables in a regression. Consequently, they conclude that the price index accounts for a much larger share of the wage disparities. More in line with my results is Lehmer and Möller (2010), who report that about half of the urban-rural wage gap is accounted for by their observables, namely firm size, the industrial composition and individual characteristics.

The relation between regional and countrywide wage inequality is recently analyzed for the US. Baum-Snow and Pavan (2013) also find that the share of high-skilled workers increases, especially in large cities. In contrast to my findings for Germany, the wage inequality between US regions increases and contributes 23% to the countrywide growth of wage dispersion. Because the regional inequality grows particularly in large cities, Baum-Snow and Pavan (2013) conjecture that agglomeration economies and demand for high-skilled workers are responsible for the regional wage differences. In instances when wages grow most where living costs are highest, and wages are deflated by regional living costs as in

Moretti (2013), the resulting utility differences are understandably less pronounced than differences in nominal wages.

Fact 2. Direct wage compensations for higher costs of living are 20%. Regional remuneration and living cost differences only account for 13% of the wage differences between regions.

Fact 3. The rise in wage inequality evolves steadily and comes almost completely from changes within counties. Wage dispersion and skill premiums grow evenly in regions of different size.

3.5 Decompositions based on the Shapley value

So far, I measured inequality by the variance and the Theil-T index. The advantage of the variance is its convenient and "natural" decomposability into an arbitrary number of components. Shorrocks (2013) provides a slightly more complicated approach that is applicable to any inequality measure. Because of its formal equivalence to the Shapley value in cooperative games, this procedure is termed *Shapley decomposition*. To evaluate the impact of a wage component W^c, an inequality index $I(\cdot)$ is first computed with the original wage w and with $w - W^c$. The difference between the two inequality indices reflects the contribution of the component W^c to the total inequality level. The contribution of the next component W^d is equal to the difference between the inequality value of $w - W^c$ and $w - W^c - W^d$. That is, the contributions S^c, S^d of both components are measured as

$$\begin{aligned} S^c &= I(w) - I(w - W^c) \\ S^d &= I(w - W^c) - I(w - W^c - W^d) \end{aligned} \tag{10}$$

This process continues until all components are sequentially eliminated. In this consideration, the decomposition is not symmetric because the value of S^c from eq. (10) differs if W^d is eliminated prior to W^c. Hence the main disadvantage of this method compared to the decomposition in section 3.3 is that the order of elimination influences the result. The solution requires that all possible elimination sequences are considered. Finally, the contribution of each component is equal to its average contribution in all of the elimination sequences. To keep the expositional clarity high and the computational burden low, I combine the wage components into three aggregate components: (a) individual-, firm- and sector-specific variables, (b) occupational attributes (as the single most important factor) and (c) regional variables.

There are two reasons why the identified contribution of components differ between inequality measures and why it is therefore advisable to compare several indices side-by-side. The Shapley decomposition yields the expected marginal impact of components on inequality. In the variance decomposition, the contribution of the components reflect their variability, cf. Kimhi (2011). The difference is that in the latter case, the contribution to inequality of an equally distributed component or a constant transfer is zero. The Gini also shares this property, but the other indices I calculate here do not, see table 5. Another

explanation rests on the fact that "different measures are underlined by different social welfare functions and are sensitive to different segments of the Lorenz curve" (Wan and Zhou 2005: 115).

It also makes a difference whether the inequality measures are calculated with the target variable in logs or in levels. The transformation into levels has the advantage that the regression constant becomes a *multiplicative* factor. Due to the scale invariance proposition of the inequality indices, the constant term has no effect on the inequality level, regardless of whether the inequality measure is invariant to constant transfers or not. The calculation in levels thus reduces the variability in the results across inequality indices and facilitates the comparison to other studies that also usually use income values in levels, compare the discussion in Morduch and Sicular (2002) and Wan (2004).

Table 5 presents the contribution of wage components in absolute and relative terms from the Shapley decomposition of the Gini, the Theil-T, the 80/20 percentile ratio and the standard deviation of logs. As expected, the explained shares of overall inequality vary between the four measures. The Theil index and the variance (see table 2) attribute higher explanatory power to the components than the remainder indices. At the other extreme, the 80/20 percentile ratio suggests that only between 19% and 28% of total wage inequality is explainable.

[Insert table 5 about here.]

Nevertheless, the components' relative changes and their ranking yield conclusions similar to the ones in section 3.3. Occupational attributes represent the most important explanatory variable and its absolute and relative contribution increases over time. Altogether, the remaining explanatory characteristics related to employers and employees account for less than the occupations' contribution in the first period, but their relative importance roughly doubles in the following years. Between 2005 and 2010, their explanatory share ranges from 12% to 26%. Again, my estimates suggest that equating local price levels and any returns due to agglomeration advantages would reduce wage inequality by only 2–5%. In the last years of the sample, their contribution even drops to zero. In sum, wage inequality increases over time but the explanatory share of observables also does. Considering the inequality indices in table 5, the observable variables are capable of explaining between 66% and 80% of the rise in wage inequality.[17] Hence the variables in this study are capable of explaining more than half of the change in raw wages.

Wan (2004) and Wan and Zhou (2005) were the first to combine the regression-based inequality decomposition with the Shapley value framework in Shorrocks (2013). Both papers examine income inequality in China and are thus not directly comparable. Nevertheless, their results also differ considerably across several decomposed inequality measures. Devicienti (2010) demonstrates the importance of the Shapley approach vis-à-vis the results

[17] For the Gini coefficient, for example, residual wage inequality in period 1 is 0.217 - 0.065 = 0.152, in period 2 it is 0.155 and in period 3 the residual wage inequality equals 0.164. The difference of 0.012 accounts for no more than 25% of the total increase of the Gini index of original wages.

from a decomposition as in Juhn *et al.* (1993), who calculate percentile ratios, but do not account for the path-dependency. Dustmann *et al.* (2009: 850) report the standard deviation of log wages in West Germany for the years 1975 to 2004. The authors use raw wages and wages adjusted for education and age. The dispersion of the former is almost equal to the one in table 5, opposed to the share of residual wages that increases from 84.2% to 87% in the period 1993–2004.

Fact 4. Different inequality indices produce strongly deviating results. The variance and the Theil index show that the wage determinants explain more than 50% of total wage inequality, whereas the 80/20 percentile ratio and the standard deviation of logs indicate that this share is about 30%. However, across all indices, between 66% and 80% of the *rise* in wage inequality are explainable and the relative importance of wage determinants is constant.

3.6 Changes in the wage distribution

So far, the considered inequality indicators shed light on which factors contribute to the rise in inequality. The next question is how different parts of the population are affected, i.e., where in the wage distribution the changes occur? Three issues will be explored in this subsection. To begin, I characterize the observed changes in original wages. I then depict the impact that each wage component has on the curvature of the wage distribution in a single year. Finally, I consider the changes in the components and how these changes add up to the total change in raw wages displayed in figure 1.

The 85/50 and the 50/15 percentile ratio of original wages in figure 3 show that workers at the top of the distribution gain relative to those in the middle, and the middle of the distribution gains relative to the bottom. Both ratios have almost the same value in 1993 but the wage gap wides by an additional 10%. The 85/50 percentile gap grows steadily by 1.5% per year until 2007 and then flattens out. Conversely, the 50/15 gap begins to rise slowly in 1998 and accelerates its growth in 2004. During the economic and financial crisis of 2007 to 2009, inequality remains relatively stable at all parts of the distribution. Concerning the residual wages $\tilde{w}_{it}$, both percentile ratios are lower and show a slow but continuous increase over time. Again, this suggests that apart from this small increase in residual wage inequality, the main drivers of the wage dispersion are captured by the factors in eq. (1). Note that for the observations in figure 3 it does not make a qualitative difference whether the 85th or the 80th percentile is used. This supports the assessment that the wage imputation, which particularly applies to wages above the 80th percentile, does not yield distorted findings.

[Insert figure 3 about here.]

The illustration in figure 3 makes clear that the rise of inequality is a continuous process. In the following, I will illustrate the changes between 1995 and 2007, which are the reference years in the middle of the first and third period for which the wage regression is estimated.

Yet there is a difference between the first and the second half of the sample period, as previewed in figure 1: Between 1995 and 2001, the first 20 percentiles suffer real wage decreases, while workers in higher percentiles experience monotonically increasing wage gains. This period is covered in Dustmann *et al.* (2009) and Antonczyk *et al.* (2010), who make the same observation. A new insight is that over the further course of the 2000s, there are considerable losses for all workers up to the 60th percentile. Even with an updated database, there is still no sign of a polarization of wages, as observed by Autor *et al.* (2008) in the US. Some reasons for the different development are identified further below.

To evaluate the impact of wage components on the curvature of the wage distribution in a single year, the additivity of the baseline regression is again exploited. I subsequently subtract component effects (as defined in eq. (4)) from the original wage and then compare the distributions with each other. The difference between the distributions of the resulting adjusted wages and the original wages provides insight regarding about the importance of the eliminated components. Because the wage regression is linear additive, there is no path dependency regarding these differentials, as opposed to the calculation of inequality indices in the preceding section. The adjusted wage distribution of, e.g. $w - X$, corresponds to a counterfactual situation in which there were no differences in unemployment and apprenticeship records prior to the current job, education, tenure, and age in the population. Since the subtraction of X is equal to setting the coefficients contained in X to zero, the counterfactual may also be interpreted as a situation in which all individuals are paid like the workers in the reference category regarding the characteristics in X.

Drawing on the prior results from aggregate inequality indices, I keep the graphical analysis of wage distributions simple and construct only three different counterfactuals. Labor income without differences in personal characteristics (red line), after additional removal of firm, regional and sectoral characteristics (green line), and after the removal of the explained part of the worker adherent skills v_{expl} (yellow line). The latter counterfactual wage is equal to the fluctuation of residual wages around the remuneration of the reference workers (i.e. the regression constant). The two graphs in figure 4 illustrate the distribution of these (adjusted) wages exemplarily for the years 1995 and 2007.[18]

[Insert figure 4 about here.]

Focus first on the comparison of the distribution in a single year. It is eye-catching that the largest part of the wage dispersion is accounted for by occupational attributes. Their share of labor income is especially large at the top of the distribution. The explanatory power of X and of $Z + K_{total} + S$ are mentionable, but both lag far behind. Although, the yellow curve of the fully adjusted wages is much flatter, it is still far from being horizontal. This is in line with the prior assessment that about 50% of the total inequality level are explainable.

[18] Note that in this and in the following figures only the 4th until the 96th percentiles are displayed to make the distribution less vulnerable to outliers. Some wages at the top are quite high and would thus distort the scale of the axis. In figure 4, the curvature of the wage distributions is very smooth, except for the spike in the upper part of the distribution which, in any case, has to be taken with care due to the censoring of wages.

The distribution expresses this observation in a more illustrative way. Original wages range from 40 € to 210 € per day in 2007, while in a world where all observable characteristics are remunerated equally, this range shrinks to 60–120 €. Comparing the change between both years, the steeper slope of the raw wage distribution in 2007 shows - from a different point of view - how inequality increases over the years.

After considering the impact of component effects in separate years, the following graphs directly show how the total wage change is composed of these components. Analyzing the distribution of these changes reveals in which direction, how strong and where exactly each of the components affects the wage dispersion in the course of time. The basic idea is as described in subsection 3.2. The change between wages is decomposed according to eq. (4) and computed separately with the average values in every percentile of the wage distribution.

$$w_{t=2007} - w_{t=1995} \equiv \Delta w = \Delta level + \Delta X + \Delta Z + \Delta S + \Delta K_{\text{total}} + \Delta v_{\text{expl}} + \Delta \xi \qquad (11)$$

In contrast to the decomposition of inequality indices, the regression constant complicates the breakdown into components. The difficulty arises because $w_{t=2007}$ and $w_{t=1995}$ are estimated in two different regressions that have two different regression constants and because the coefficients of all categorical variables need to be interpreted relative to their respective regression constant. S, for example, reports the average payment differential in sectors relative to the reference sector. Likewise, ΔS, reports the wage change in every sector relative to the change in the reference sector, which is embodied in $\Delta level$. The point is that the regression constant ($level$) represents the reference for *all* categorical variables, but cannot be apportioned to the different categorical components. Fortin *et al.* (2011: 45) discuss this issue in depth but the conclusion is that "there is no quick fix", especially in the present case with a multitude of categorical variables and interactions between them. The good news is that the interest here is not in a single coefficient but on the overall effect of aggregate components. And even if the exact level of component effects is meaningless, its contribution to the overall change in inequality can be inferred from figure 5. Departing from $\Delta level$, I subsequently add changes of component effects until the total change of raw wages is obtained. The direction of the induced shift of the curve gives a clear statement about the contribution of the added component.

The wage change of the reference worker is about +8.5%, depicted by the horizontal line in figure 5. Note that the order in which components are added does not influence the result and is only chosen for expositional purpose. Adding the change in firm-specific payments shifts the lower end of the distribution about 5 percentage points down and helps to explain some of the gains of the top 40% of earners. Thus, these factors clearly contribute to the rise in wage inequality. Once the change in sectoral characteristics is added, the curve is shifted down but more so at the left end. This means that all workers, on average, gain less than those workers in the reference sector (car repairing). A closer inspection of individual sectors reveals that the negative change is particularly due to the public sector and the service sector. Adding the change in X does not influence workers in the fourth

quartile much, whereas wages of workers further left in the distribution develop more and more unfavorably. The interpretation for this finding is intuitive. The reference worker has medium education (which applies to more than 70% of the population, cf. table 1), and most likely shares some of 8.5% wage increase in $\Delta level$. Since the changes in the upper part of the distribution are close to zero, the average payments to highly educated workers are minor. On the other side, the payment to the low-skilled shrinks. This group is spread across the entire wage distribution, but is naturally more concentrated at the bottom, which explains the unequal downward shift after adding X. Region-specific payments altogether show a negative effect in the upper half of the distribution but show virtually no effect at the bottom half. This issue is discussed at length in subsection 3.4 where we identified that the average size of firms in agglomerations shrinks and that firm-specific wage premiums in agglomerations decreases relative to the payments in rural regions (the reference category). Thus, figure 5 shows that evidently employees of large firms in agglomerations are mainly found in the upper half of the wage distribution.

[Insert figure 5 about here.]

In the lower graph in figure 5, the black line is generated by augmenting the prior change with the change attributed to the workers' occupational attributes. Like in the exercises above, this explained part of the fixed effect turns out to be the most important wage component. In contrast to the remaining characteristics, it accounts for substantial gains of high-wage workers. Finally, it is remarkable that the remaining unexplained parts of the wage regression deviate around zero, i.e., the black line is already very close to the overall wage change. Therefore, the wage residual does not have a significant impact on the increase of the wage dispersion. Altogether, these observations naturally agree with the prior, albeit less detailed, conclusions from the aggregate inequality measures. I identify the sources of changes in those wage components more in-depth in the next subsection.

Fact 5. Changes related to occupations have a particularly large impact on the top of the wage distribution. Changes in the service sector contribute to wage decreases of low-wage workers. None of the considered wage determinants contribute to a polarization of wages. Except for region-specific factors, all determinants add to the decrease the lower end of the wage distribution.

3.7 Endowment and price effects

In the previous subsection, I divide the total wage change into the impact of its components in every percentile of the distribution. Although this decomposition reveals in which part of the distribution inequality decreases or increases, it does not provide a particular reason why. For example, v_{expl} may shift the wage distribution upwards because more workers have high-wage occupation and/or because high-wage occupations have a better remuneration. In the following, I subdivide the impacts of components into fractions attributed to their composition and coefficients. This is the basic idea in the seminal papers

by Blinder (1973) and Juhn *et al.* (1993) that is described in section 3.2. I now apply the decomposition in eq. (3) to the separate components in eq. (11). Thereby, the total wage change may be re-arranged so that

$$\begin{aligned}
&\Delta w_i \equiv w_{i,t=2007} - w_{i,t=1995} && (12)\\
&= \Delta\beta_0 + \Delta\beta_x X_{i,t=2007} + \Delta\beta_z Z_{j,t=2007} + ... + ... && \text{... coefficient effect}\\
&+ \beta_{x,t=1995}\Delta X_i + \beta_{z,t=1995}\Delta Z_j + ... + ... && \text{... composition effect}\\
&+ \Delta\upsilon_{unexpl} + \Delta\epsilon_i && \text{... residual effect}
\end{aligned}$$

As in the prior decomposition, I calculate these effects for each percentile of the wage distribution. Figure 6 presents the composition, coefficient and residual effect in aggregate. As can be seen, they amount exactly to the total wage change. For a start, it is more convenient to show aggregate changes because only for personal, employer and sectoral attributes is the composition change, evaluated at the return to these characteristics in 1995, in some parts of the distribution larger than 1%. Note that for the qualitative shape of the curves in figure 6 it is not decisive whether coefficient changes are evaluated by the value of the characteristics in 2007 or the other way round. The same is true for the illustration of the composition effect. In any case, the coefficient effect evidently dwarfs both other effects, so that differing returns mainly account for the wage change over the past 20 years.[19] This findings suggest that not supply but rather demand or performance-based explanations are responsible for the pronounced increase in wage inequality.

[Insert figure 6 about here.]

By drawing on the number and remuneration of high-skilled workers in different regions in table 3, we have already understood the reasons for the change in regional characteristics sufficiently well. Next, I break down the shifts of X, Z and υ both into their separate variables that these components combine, and into the detailed effects according to eq. (12). Regarding Z, the main effect can be traced back to the size of the plant. Table 2 shows that the covariance between Z and the wage is positive and increases strongly between 1995 and 2007. In line, the coefficients in the wage regression indicate that there is a positive firm size premium. The right graph in figure 7 reveals a pronounced assortative matching of high-skilled workers and large firms and matching between low-wage workers and small firms. Because the average firm size shrinks for workers below the 75th percentile during the considered period, the sorting of workers becomes even more pronounced. Nevertheless, the left graph of figure 7 shows that this composition change is minor compared to the change in returns to firm size.[20] Obviously, the compensation differential between small and large firms grows, so that low-wage workers in small firms earn even less and the other way

[19] It should again be noted that this decomposition method assumes that workers in all parts of the wage distribution receive the same return to their characteristics. Because the remaining wage residual is close to zero, this assumption does not seem to be unduly wrong.

[20] The consideration of the component and the coefficient effect already incorporates the diverging returns in East and West Germany and how both of them changed.

round. Comparing the range of -4% to +2% of this change to the total component effect of Z in figure 5 confirms that the firm size effect is the main driver of this development.

[Insert figure 7 about here.]

In the same way, ΔX can be partitioned into its variables. Three of them are principally responsible for the change related to individual characteristics: age, unemployment prior to the current job and the lump-sum payment differences between workers in East and West Germany. According to Eurostat data, in the period between 1995 and 2005 the German unemployment rate increases from about 8% to 11%. Summary Statistics in table 1 confirm that 8% of the workers in the sample have an unemployment record. The identified wage cut that these workers face when re-entering the labor market changes from -5.5% to -7.4%. Because this variable is either 0 or 1, the coefficient effect in figure 8 clearly reflects the monotonic decreasing probability of becoming unemployed along the wage distribution. The composition effect shows that the average change for high-wage workers is close to zero, whereas workers in the lower half of the wage distribution are more likely to have an unemployment record. In this case, both the coefficient and the composition effect contribute to the rising wage inequality.

The right graph in figure 8 displays the same breakdown for the workers' age structure. It is important to note that the return to age is positive and basically increasing. The average age among high-wage workers hardly changes and thus the composition effect is only slightly positive. However, between the 15th and the 60th percentile the age structure induces wage gains of 2–3%. Nevertheless, a separate consideration of the average age (analog to plant size in figure 7) shows a steady increase along the wage distribution and an age gap of 10 years in both 1995 and 2007. The coefficient effect of age in figure 8 shows a similar slope and thus reflects the fact that the returns to experience in terms of age rise. Finally, the coefficients of the West dummy in the first and last period indicate that the average wage gap falls by 4 percentage points. Yet, controlling for observable characteristics and their diverging returns in East and West, the East German workers still receive 6.7% less than their West German counterparts.

[Insert figure 8 about here.]

The largest share of the aggregate composition and coefficient change is explained by occupation specific attributes. These are already noteworthy in themselves because they are rarely the focus of studies about income inequality. Occupations are either not considered at all, or their pivotal impact is missed because they are not related to workers' unobserved or adherent skills. Recall that the estimated worker fixed effects from the wage regression are used in a second step and are regressed against 3-digit occupation dummies. The explained part in the latter regression thus represents the portion of the countrywide wage dispersion, which may be traced back to specific requirements and abilities in occupations. As before, it can be seen in figure 9 that the change in returns is much larger and that

it benefits workers in the upper half of the wage distribution but disadvantages the lower half. Furthermore, there are some noteworthy gains due to a higher stock of high-skilled occupations.

[Insert figure 9 about here.]

To subdivide and depict changes of these 341 different occupations, a combination introduced by Matthes *et al.* (2008) is used. Based on data from the Federal Employment Agency about required skills and typical tasks in occupations, the authors calculate degrees of similarity and classify similar occupations into 21 "segments".[21] Interestingly, the development of segments is highly heterogeneous and exhibits different patterns. The further consideration is restricted to 10 of the 21 segments where wage changes exceed 1%. For convenience, either two or three segments, where the patterns resemble each other, are pooled in figure 10. Still, the arrangement along the wage distribution allows for differentiation between occupations in the same segment.

Consider the pattern of jobs related to metal processing and construction in panel (d). In the right part of the distribution, engineers, architects, etc. are located. Unskilled and blue collar workers are represented at the other end. The latter's supply clearly increases over time while the supply of high-skilled workers diminishes. The price for these types of labor moves in the opposite direction of the employment change. Consequentially, the development seems to be driven by the labor supply. This is not the case in the other segments in panels (a) to (c). Jobs disappear and at the same time are paid less in the gastronomy, logistics and the residual segment (mainly comprising laborers without a detailed job description). On the other hand, jobs in the electronics industry, as well as managers, economists, lawyers, doctors and engineers (to give some example of prosperous occupations in the segments shown in panels (b) and (c)) exhibit a higher employment share and higher remuneration.

What could be the reason for these developments, given that observable differences in personal, sectoral and firm-specific attributes between occupations are already taken into account in the wage regression? One possible explanation is that relative productivity growth and resulting payment increases are disproportionately distributed across occupations. Another explanation is that the rise in demand for some occupations and skills exceeds the available supply. Both possibilities are consistent with the skill-biased technical change that is accompanied by the computerization of the economy (Autor *et al.* 2003). This development overly increases the demand for analytical and interactive skills. Spitz-Oener (2006) confirms that these skills are used intensively by employees with a high level of education, i.e., by individuals at the upper end of the wage distribution. Autor *et al.* (2008) also find that the rise of the wage inequality in the US features higher employment and higher wages of the most skilled workers; however, they do not link these changes to

[21] The information in their data is applied by the Employment Agencies to propose similar jobs to unemployed workers and suitable job candidates to firms who failed to match with their current occupation specification. Therefore, Matthes *et al.* (2008) argue that their combination into segments aligns similar occupations better than the official 1- or 2-digit classification.

occupations and skills as in the present approach. The US also experience a wage growth at the left end of the wage distribution relative to its middle, where jobs involve many routine tasks which are negatively affected by SBTC. This development is much weaker in Europe, as already noted by Acemoglu and Autor (2011) and Antonczyk *et al.* (2010), because of the increased supply of low-skilled workers and deunionization.

[Insert figure 10 about here.]

Fact 6. Differing returns mainly account for the wage change over the past 20 years. Low-wage workers are more likely to be young, work in small firms and become unemployed. The relative returns to these characteristics even decrease. Higher supply and higher remuneration in occupations related to management and engineering are responsible for a large part of the widening gap at the top of the wage distribution.

4 Conclusion

The present study complements the knowledge about the rise in wage inequality over the last 20 years in Germany. Inequality increases more in the 2000s than in the 1990s, whereby workers below the 60th percentile see their real wage decline. The overall change in the wage distribution is monotonically increasing from -15% at the left end to +20% at the right end. As most of the preceding studies conclude, an ever smaller share of the rising wage inequality is explained by education, experience and and industry-specific variables. I dismantle this residual wage further by additionally considering workers' occupations, characteristics of their employers, the local price level as well as regional productivity and payment differences. Altogether, more than 50% of the wage dispersion and more than 75% of the increase in wage inequality are explained by this approach.

I apply regression-based decomposition methods that allow the quantification of the separate impact of these wage determinants on aggregate inequality indices and on the entire wage distribution. It turns out that occupational attributes are the most important wage determinants. Highly paid technical, administrative and business occupations in particular have higher returns, even though the supply in these occupations increased. In general, the data show that the changes in the composition workers' characteristics play a minor role compared to changes in returns. Both findings may be explained by skill-biased technical change. Many prior studies argue that SBTC is the main driver for the rise in wage inequality, however, the conclusion in this paper is not merely based on the residual wage. In the lower percentiles of the wage distribution, particularly workers in the service, construction and logistics sectors suffer real wage losses. This finding is in line with the higher supply of low-skilled workers which explains why Germany does not experience a wage polarization.

Another reason for gains in the upper part of the wage distribution and declines in its lower part is the assortative matching between skilled workers and large firms, and particularly the increase in the firm size premium. In addition, workers in the lower part of the

distribution are on average younger and thus do not benefit from increased returns to experience. They are also more likely to have an unemployment record for which wage discounts are magnified. Even though high-skilled workers and large firms tend to be concentrated in large regions and even though there is a skill premium in larger regions, these differentials do not intensify over time. Therefore, the rise in wage inequality is distributed equally across regions and consequently coming from changes within counties. All in all, regional price levels, wage premiums and diverging returns in regions of different size have a low explanatory power of wages that even decreases over time.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4, Amsterdam: Elsevier-North Holland, pp. 1043–1171.

Andersson, F., Burgess, S. and Lane, J. (2007). Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics*, **61** (1), 112–128.

Antonczyk, D., DeLeire, T. and Fitzenberger, B. (2010). Polarization and rising wage inequality: comparing the US and Germany. *IZA Discussion Paper No. 4842*.

Autor, D., Dorn, D. and Hanson, G. H. (2013). The China Syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, **103** (6), 2121–2168.

—, 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.

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.

Baldwin, J. R., Brown, W. M. and Rigby, D. L. (2010). Agglomeration economies: Microdata panel estimates from Canadian manufacturing. *Journal of Regional Science*, **50** (5), 915–934.

Baum-Snow, N. and Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, **95** (5), 1535–1548.

Baumgarten, D., Geishecker, I. and Görg, H. (2013). Offshoring, tasks, and the skill-wage pattern. *European Economic Review*, **61**, 132–152.

Berman, E., Bound, J. and Machin, S. (1998). Implications of skill-biased technological change: international evidence. *Quarterly Journal of Economics*, **113** (4), 1245–1279.

Biewen, M. and Juhasz, A. (2012). Understanding rising income inequality in Germany, 1999/2000–2005/2006. *Review of Income and Wealth*, **58** (4), 622–647.

Blien, U., Gartner, H., Stüber, H. and Wolf, K. (2009). Regional price levels and the agglomeration wage differential in western Germany. *The Annals of Regional Science*, **43** (1), 71–88.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, **8** (4), 436–455.

Brown, C. and Medoff, J. (1989). The employer size-wage effect. *Journal of Political Economy*, **97** (5), 1027–1059.

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

Combes, P., Duranton, G. and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, **63** (2), 723–742.

Dauth, W., Findeisen, S. and Südekum, J. (2014). The rise of the East and the Far East: German labor markets and trade integration. *Journal of the European Economic Association*, **forthcoming**.

Devicienti, F. (2010). Shapley-value decompositions of changes in wage distributions: a note. *Journal of Economic Inequality*, **8** (1), 35–45.

DiNardo, J., Fortin, N. M. and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, **64** (5), 1001–1044.

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

Ebenstein, A., Harrison, A., McMillan, M. and Phillips, S. (2014). Estimating the impact of trade and offshoring on american workers using the current population surveys. *Review of Economics and Statistics*, **forthcoming**.

Eckey, H.-F., Kosfeld, R. and Türck, M. (2006). Abgrenzung deutscher arbeitsmarkt-tregionen. *Raumforschung und Raumordnung*, **64** (4), 299–309.

Ehrl, P. (2013). Agglomeration economies with consistent productivity estimates. *Regional Science and Urban Economics*, **43** (5), 751–763.

— (2014). High-wage workers and high-productivity firms - a regional view on matching in germany. *mimeo*.

Fitzenberger, B., Osikominu, A. and Völter, R. (2006). Imputation rules to improve the education variable in the IAB employment subsample. *Schmollers Jahrbuch*, **126** (3), 405–436.

Fortin, N., Lemieux, T. and Firpo, S. (2011). Decomposition methods in economics. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4, Part A, *1*, Elsevier, pp. 1–102.

Foster, J. E. and Ok, E. A. (1999). Lorenz dominance and the variance of logarithms. *Econometrica*, **67** (4), 901–907.

Gartner, H. (2005). The imputation of wages above the contribution limit with the German IAB employment sample. *FDZ Methodenreport*, **2/2005**.

Gibbons, R. and Katz, L. (1992). Does unmeasured ability explain inter-industry wage differentials? *Review of Economic Studies*, **59** (3), 515–535.

Gibbons, S., Overman, H. G. and Pelkonen, P. (2013). Area disparities in Britain: Understanding the contribution of people vs. place through variance decompositions. *Oxford Bulletin of Economics and Statistics*.

Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany. *Journal of Labor Economics*, **30** (1), 175–213.

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

Helpman, E., Itskhoki, O., Muendler, M.-A. and Redding, S. J. (2014). *Trade and inequality: From theory to estimation*. Tech. rep., mimeo.

Idson, T. L. and Oi, W. Y. (1999). Workers are more productive in large firms. *American Economic Review: Papers and proceedings*, **89** (2), 104–108.

Juhn, C., Murphy, K. M. and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, **101** (3), 410–442.

Kawka, R., Beisswenger, S., Costa, G., Kemmerling, H., Müller, S., Pütz, T., Schmidt, H., Schmidt, S. and Trimborn, M. (2009). *Regionaler Preisindex*. Berichte Band 30, Bundesinstitut für Bau-, Stadt- und Raumforschung, Bonn.

Kimhi, A. (2011). Comment: On the interpretation (and misinterpretation) of inequality decompositions by income sources. *World Development*, **39** (10), 1888–1890.

Kosfeld, R. and Eckey, H. (2010). Market access, regional price level and wage disparities: the German case. *Jahrbuch für Regionalwissenschaft / Review of Regional Research*, **30** (2), 105–128.

Krueger, A. B. and Summers, L. H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica*, **56** (2), 259–93.

Lehmer, F. and Möller, J. (2010). Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for germany. *The Annals of Regional Science*, **45** (1), 31–53.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, **96** (3), 461–498.

Lindley, J. and Machin, S. (2013). Wage inequality in the labour years. *Oxford Review of Economic Policy*, **29** (1), 165–177.

— and — (2014). Spatial changes in labour market inequality. *Journal of Urban Economics*, **79**, 121–138.

Matthes, B., Burkert, C. and Biersack, W. (2008). Berufssegmente: Eine empirisch fundierte neuabgrenzung vergleichbarer beruflicher einheiten. IAB Discussion Paper.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, **71** (6), 1695–1725.

Morduch, J. and Sicular, T. (2002). Rethinking inequality decomposition, with evidence from rural China. *Economic Journal*, **112** (476), 93–106.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, **5** (1), 65–103.

Mortensen, D. T. (1987). Job search and labor market analysis. In O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, *Handbook of Labor Economics*, vol. 2, *15*, Elsevier, pp. 849–919.

Puga, D. (2010). The Magnitude and Causes of Agglomeration Economies. *Journal of Regional Science*, **50** (1), 203–219.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, **90** (6), 1257–1278.

Shorrocks, A. F. (1982). Income decomposition by factor components. *Econometrica*, **50** (1), 193–212.

— (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, **11** (1), 99–126.

— and Wan, G. (2005). Spatial decomposition of inequality. *Journal of Economic Geography*, **5** (1), 59–81.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, **24** (2), 235–270.

Suedekum, J. (2006). Agglomeration and regional costs of living. *Journal of Regional Science*, **46** (3), 529–543.

Van Reenen, J. (2011). Wage inequality, technology and trade: 21st century evidence. *Labour economics*, **18** (6), 730–741.

vom Berge, P., König, M. and Seth, S. (2013). Sample of integrated labour market biographies (siab) 1975-2010. *FDZ Datenreport*, **1/2013**.

Wan, G. (2004). Accounting for income inequality in rural China: a regression-based approach. *Journal of Comparative Economics*, **32** (2), 348–363.

— and Zhou, Z. (2005). Income inequality in rural China: Regression-based decomposition using household data. *Review of Development Economics*, **9** (1), 107–120.

Winters, J. V. (2009). Wages and prices: Are workers fully compensated for cost of living differences? *Regional Science and Urban Economics*, **39** (5), 632–643.

Data Appendix

Around 12% of wages are top-coded in the SIAB data due to a upper limit of the contribution to the pension fund. Because wages are quite different in East and West Germany, this limit is also different in each year between both parts. In the joint wage distribution, the East German upper bound lies at the 71th percentile in 1993, at the 78th percentile in 1994 and oscillates afterward between the 81th and the 85th percentile. For some of the inequality measures such as the 80/20 percentile ratio, it is possible to avoid the dependence on imputed wages. However, in the figures with (adjusted) wage distributions, the part above the 81th percentile show some discontinuities and thus need not be taken at face value. Nevertheless, it is necessary and common to rehabilitate the entire wage distribution by employing an imputation procedure for the top-coded wages based on censored regressions proposed by Gartner (2005). This imputation is also applied in the analysis of wage inequality by Dustmann *et al.* (2009) and Card *et al.* (2013). I estimate a Tobit model and compute the censored wages as the sum of two components: The predicted wage plus a random draw from a truncated normal distribution (the distribution of log wages). The Tobit estimation includes third order polynomials of age and tenure, the plant's size and shares of high- and low-skilled workers, and dummies for education, occupation groups, industries, years, location in West Germany, German citizenship and unemployment prior to the current job. A different imputation procedure where education groups, years and East and West Germany are estimated separately using occupation segment dummies, the plant's size and shares of high and low-skilled workers and second order polynomials of tenure and age basically yields the same results.[22] Afterward, the imputed wages are deflated with the national consumer price index (CPI), where 2005 is the base year.

Another imputation is necessary for the education variable because it exhibits about 10% missing values and 0.4% temporal inconsistencies, see also Dustmann *et al.* (2009) and Fitzenberger *et al.* (2006). These values are replaced using the panel dimension of the data set. For example, if an individual has a medium or high education level in all but one year, and this year is not equal to the first observation, this inconsistent education 'downgrade' is replaced by the prior value. If the person exhibits more than one missing value, the imputation is only applied if the valid values before and after the missing are the same, or if the gap is in the beginning or the end of the period. Other cases are not encoded to avoid overconfident predictions. Altogether, three education categories are distinguished: low (high-school equivalent), medium (college equivalent and/or vocational training) and high (university graduates).

[22] Both Dustmann *et al.* (2009) and Card *et al.* (2013) also conduct several robustness checks regarding the imputation procedure and obtain very similar results regarding the residual wage inequality.

For the construction of the final sample, observations with missing values in any variable are disregarded. This study only considers full-time employment relations where the wage is above the official marginal part-time income threshold. A data-driven justification for this restriction is that the SIAB does only provide the daily wage but not the hours worked. Therefore, it is not possible to infer a hourly or full-time equivalent wage for part time workers. In rare cases where the data indicates that workers have multiple (full-time) jobs, merely the one with the highest wage is regarded. Remuneration in so-called "mini-jobs" is determined by politics and several exceptions regarding the social security and tax contribution imply that the wage determination cannot be compared to regular jobs. Due to these restrictions the overall extent of inequality is underestimated. The sample is further reduced to male workers between the age of 17 and 64. Finally, all individuals with less than 2 observations in each of the six-year periods are dropped because of the panel estimations. The following table 1 provides the mean and the standard deviation of wages before and after imputation as well as other variables used in the estimation of the wage regression in each of the three periods.

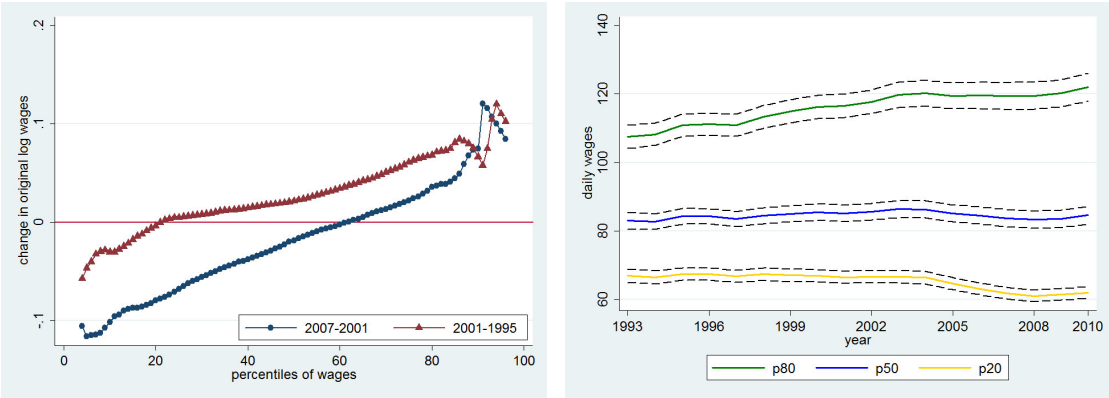
Table 1: Summary statistics

	1993–1998		1999–2004		2005–2010	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
log wage	4.48	0.37	4.51	0.40	4.49	0.44
imputed log wage	4.50	0.40	4.53	0.44	4.52	0.49
tenure (days)	2778	2456	2935	2699	3212	2908
age	39.8	10.7	40.6	10.1	41.8	10.2
prior apprenticeship dummy	0.08	0.27	0.07	0.26	0.08	0.27
prior unemployment dummy	0.22	0.41	0.27	0.44	0.30	0.46
plant size	1352	5024	1227	4833	1160	4597
plant low-qual. share	0.18	0.19	0.16	0.18	0.13	0.16
plant high-qual. share	0.08	0.14	0.09	0.15	0.10	0.16
log price index	-0.07	0.07	-0.07	0.07	-0.07	0.07
West dummy	0.80	0.40	0.82	0.39	0.83	0.38
low-education (share)	0.14	0.35	0.13	0.34	0.12	0.32
medium-education (share)	0.76	0.43	0.75	0.43	0.75	0.44
high-education (share)	0.10	0.30	0.12	0.33	0.14	0.35
county type 1 (share)	0.29	0.45	0.28	0.45	0.28	0.45
county type 2 (share)	0.16	0.37	0.17	0.37	0.17	0.37
county type 3 (share)	0.09	0.29	0.09	0.29	0.09	0.29
county type 4 (share)	0.08	0.27	0.08	0.27	0.08	0.27
county type 5 (share)	0.17	0.38	0.18	0.38	0.18	0.38
county type 6 (share)	0.21	0.41	0.20	0.40	0.21	0.40
observations	1,618,650		1,472,890		1,372,059	

Notes: Mean and standard errors for the variables necessary to estimate the wage regression in eq. (1). The summary statistics and the entire sample is divided into the three periods for which the regression is estimated. The last column shows the number of observations in each period. *Log wage* is the top-coded wage as given in the data. *Education* and *county type* are categorical variables for which the share in each category is reported. These and the remainder variables are as described in the main text.

Tables and figures

Figure 1: The development of wage inequality in Germany

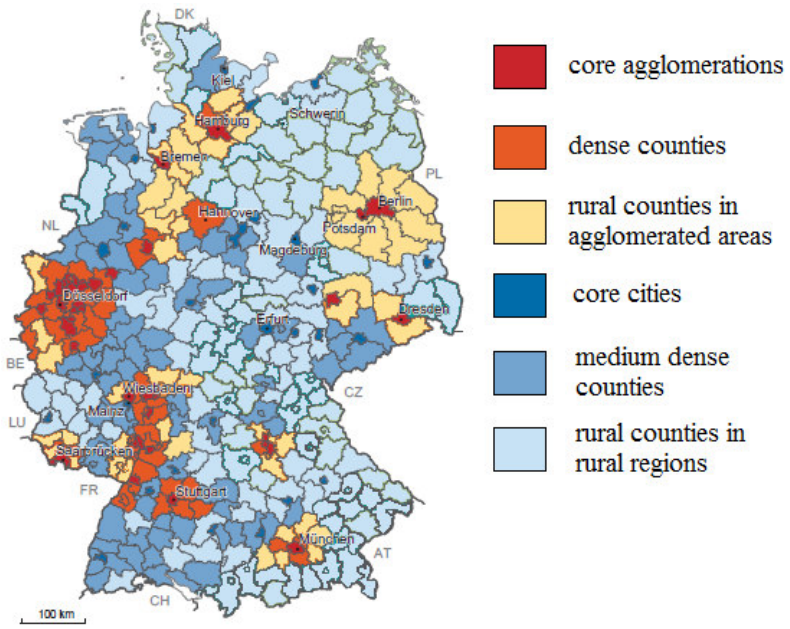


(a) Distribution of changes in log wages between 1995–2001 and between 2001–2007

(b) Wage percentiles in labor market regions

Notes: Both graphs are constructed using the SIAB data set and its imputed wages, as described in section 2. The left graph displays the change in log wages in each percentile (between the 4th and the 96th). The construction of the right graph uses the 80th, 50th and 20th percentile in each labor market region. These labor market regions are a combination of counties, as described in Eckey *et al.* (2006). The solid lines show the mean of each of the three percentiles across all labor market regions and the dashed lines mark confidence intervals around these mean values.

Figure 2: County types in a map of Germany



Notes: The figure is taken from the BBSR and has been modified slightly to aggregate the counties to six different types. Agglomerations have more than 300.000 inhabitants. Dense counties are specified to have a density of more than 300 inhabitants per square kilometer. Core cities have between 300.000 and 100.000 inhabitants. Medium-dense counties have a density of more than 150 inhabitants per square kilometer. Rural regions are less dense than 150 inhabitants per km^2 and are not close to any agglomeration.

Table 2: Variance decomposition

	total covariance		between counties		within counties	
			Panel 1: 1993–1998			
w	0.160	100.0%	0.032	20.2%	0.128	79.8%
X	0.014	8.5%	0.008	5.0%	0.006	3.5%
Z	0.002	1.2%	0.000	-0.1%	0.002	1.3%
XK	0.001	0.7%	0.000	0.0%	0.001	0.6%
ZK	0.002	1.1%	0.001	0.9%	0.001	0.3%
K	0.001	0.7%	0.001	0.7%	0.000	0.0%
Pi	0.002	1.1%	0.002	1.1%	0.000	0.0%
T	0.000	0.2%	0.000	0.0%	0.000	0.2%
S	0.005	3.2%	0.001	0.4%	0.005	2.9%
υ_{expl}	0.046	28.5%	0.006	3.6%	0.040	24.9%
υ_{res}	0.074	46.4%	0.014	8.7%	0.060	37.7%
ϵ	0.014	8.5%	0.000	0.0%	0.014	8.5%
			Panel 2: 1999–2004			
w	0.195	100.0%	0.030	15.6%	0.165	84.4%
X	0.034	17.4%	0.007	3.5%	0.027	13.9%
Z	0.008	4.3%	0.003	1.4%	0.006	2.9%
XK	0.001	0.3%	-0.001	-0.3%	0.001	0.5%
ZK	0.001	0.5%	0.001	0.6%	0.000	0.0%
K	0.001	0.2%	0.001	0.2%	0.000	0.0%
Pi	0.002	1.0%	0.002	1.0%	0.000	0.0%
T	0.000	0.0%	0.000	0.0%	0.000	0.0%
S	0.007	3.7%	0.001	0.3%	0.007	3.4%
υ_{expl}	0.055	28.0%	0.007	3.6%	0.048	24.5%
υ_{res}	0.072	36.9%	0.010	5.2%	0.062	31.7%
ϵ	0.015	7.6%	0.000	0.0%	0.015	7.6%
			Panel 3: 2005–2010			
w	0.238	100.0%	0.033	14.0%	0.205	86.0%
X	0.028	11.7%	0.005	2.0%	0.023	9.7%
Z	0.017	6.9%	0.006	2.6%	0.010	4.3%
XK	0.001	0.4%	0.000	-0.2%	0.001	0.6%
ZK	-0.004	-1.5%	-0.001	-0.5%	-0.003	-1.1%
K	0.001	0.2%	0.001	0.2%	0.000	0.0%
Pi	0.002	0.7%	0.002	0.7%	0.000	0.0%
T	0.000	0.0%	0.000	0.0%	0.000	0.0%
S	0.008	3.4%	0.000	0.0%	0.008	3.4%
υ_{expl}	0.076	32.0%	0.009	3.9%	0.067	28.2%
υ_{res}	0.096	40.3%	0.013	5.3%	0.084	35.1%
ϵ	0.014	5.9%	0.000	0.0%	0.014	5.8%

Notes: Each panel shows the results from a variance decomposition of the estimated component effects identified in the wage regression in eq. (1). Component effects are defined as the estimated coefficient multiplied by the variable. For example, the sectoral impact on the wage is $S \equiv \hat{\beta}_s D_s$. For the other definitions see eq. (4) and (5). The values in the first and second column represent the absolute and relative amounts of components from the aggregate variance decomposition according to eq. (6). Columns 3 to 6 display the division of these values into the dispersion between and within counties according to eq. (8).

Table 3: Regional summary statistics

		period 1: 1993–1998		period 3: 2005–2010		
county-type	variable	mean	ratio to overall mean	mean	ratio to overall mean	change
1	firm size	2337	173%	1621	140%	-31%
2	firm size	1427	106%	1183	102%	-17%
3	firm size	339	25%	374	32%	10%
4	firm size	2975	220%	3560	307%	20%
5	firm size	530	39%	552	48%	4%
6	firm size	398	29%	448	39%	13%
1	high-skilled share	0.15	149%	0.21	152%	42%
2	high-skilled share	0.09	92%	0.15	105%	58%
3	high-skilled share	0.08	81%	0.10	70%	19%
4	high-skilled share	0.13	125%	0.17	125%	38%
5	high-skilled share	0.07	68%	0.09	68%	38%
6	high-skilled share	0.06	63%	0.08	56%	22%
1	high-skill premium	1.15	100.66%	1.19	100.85%	3.63%
2	high-skill premium	1.14	99.69%	1.18	99.60%	3.33%
3	high-skill premium	1.14	100.15%	1.17	99.32%	2.57%
4	high-skill premium	1.14	100.05%	1.20	101.55%	4.98%
5	high-skill premium	1.14	99.55%	1.17	99.28%	3.15%
6	high-skill premium	1.14	99.61%	1.18	99.48%	3.29%
1	log wage	4.60	102.43%	4.63	102.55%	0.58%
2	log wage	4.59	102.13%	4.61	102.08%	0.41%
3	log wage	4.40	97.85%	4.40	97.43%	0.03%
4	log wage	4.49	99.84%	4.54	100.42%	1.04%
5	log wage	4.46	99.19%	4.47	98.95%	0.22%
6	log wage	4.34	96.61%	4.37	96.66%	0.51%

Notes: The definition of the six different county types is made in figure 2. High skilled individuals are defined as university graduates and the high skilled premium is calculated with respect to workers with the lowest educational attainment.

Table 4: Theil decomposition

		1993	1994	1995	1996	1997	1998	1999	2000	2001
w_{it}	wthn.	0.059	0.062	0.064	0.066	0.067	0.069	0.074	0.077	0.079
	btw.	0.018	0.016	0.015	0.015	0.015	0.015	0.015	0.015	0.015
$\tilde{w}_{it}$	wthn.	0.032	0.034	0.034	0.035	0.036	0.037	0.036	0.038	0.039
	btw.	0.004	0.004	0.003	0.003	0.003	0.003	0.002	0.002	0.002

		2002	2003	2004	2005	2006	2007	2008	2009	2010
w_{it}	wthn.	0.080	0.083	0.085	0.090	0.095	0.099	0.103	0.102	0.104
	btw.	0.015	0.016	0.016	0.016	0.017	0.017	0.018	0.018	0.018
$\tilde{w}_{it}$	wthn.	0.039	0.039	0.040	0.039	0.041	0.043	0.044	0.045	0.046
	btw.	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.003

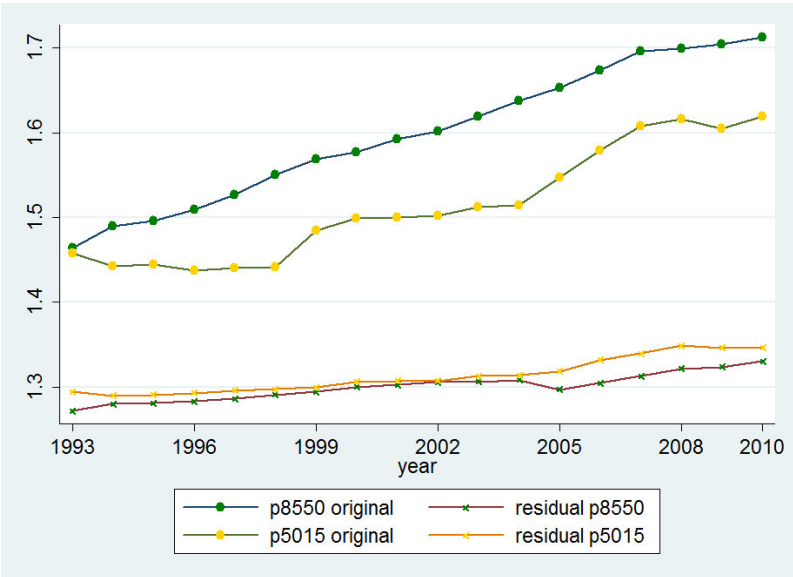
Notes: Theil indices decomposed into within and between county parts for the imputed wage in levels and the residual wage (also transformed into levels) in every year between 1993 and 2010.

Table 5: Shapley decomposition

	Gini		Theil		p8020		sd. of logs	
				Panel A: 1993–1998				
w	0.217	100%	0.080	100%	1.841	100%	0.400	100%
X+Z+T+S	0.017	8%	0.010	12%	0.099	5%	0.036	9%
υ_{expl}	0.044	20%	0.028	35%	0.216	12%	0.067	17%
K+Pi+XK+ZK	0.007	3%	0.004	5%	0.033	2%	0.011	3%
				Panel B: 1999–2004				
w	0.238	100%	0.095	100%	1.982	100%	0.442	100%
X+Z+T+S	0.032	14%	0.021	23%	0.174	9%	0.062	14%
υ_{expl}	0.048	20%	0.031	33%	0.265	13%	0.074	17%
K+Pi+XK+ZK	0.003	1%	0.002	2%	0.014	1%	0.005	1%
				Panel C: 2005–2010				
w	0.266	100%	0.116	100%	2.189	100%	0.488	100%
X+Z+T+S	0.042	16%	0.030	26%	0.270	12%	0.078	16%
υ_{expl}	0.061	23%	0.042	36%	0.356	16%	0.095	20%
K+Pi+XK+ZK	-0.001	0%	-0.001	-1%	-0.002	0%	-0.001	0%

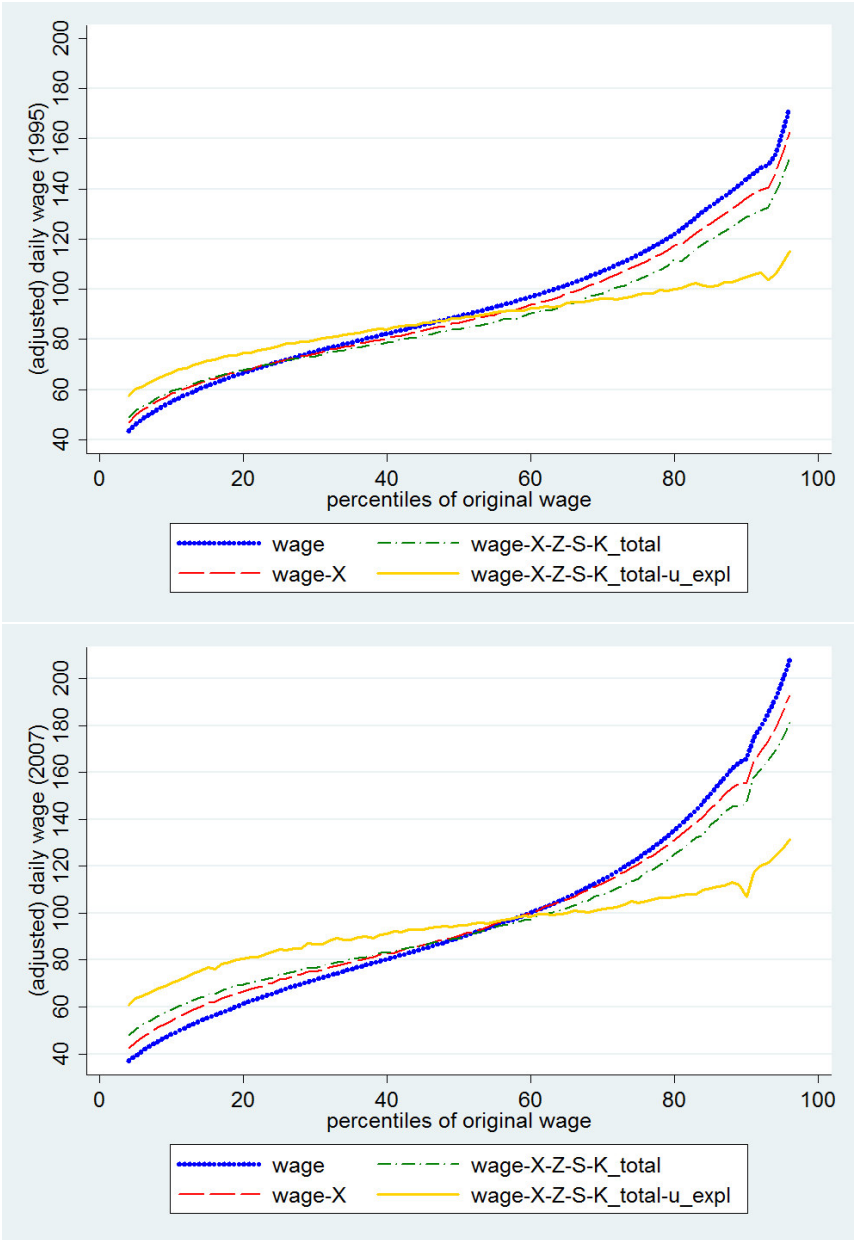
Notes: Shapley decompositions of the Gini index, the Theil index the 80/20 percentile ratio and the standard deviation of log wages. The first row reports the inequality value for the raw wage. Rows 2–4 indicate the importance of aggregate components, as defined in eq. (4). For the calculation of these inequality measures, the original and adjusted wages are transformed into levels.

Figure 3: Percentile ratios of original and residual wages



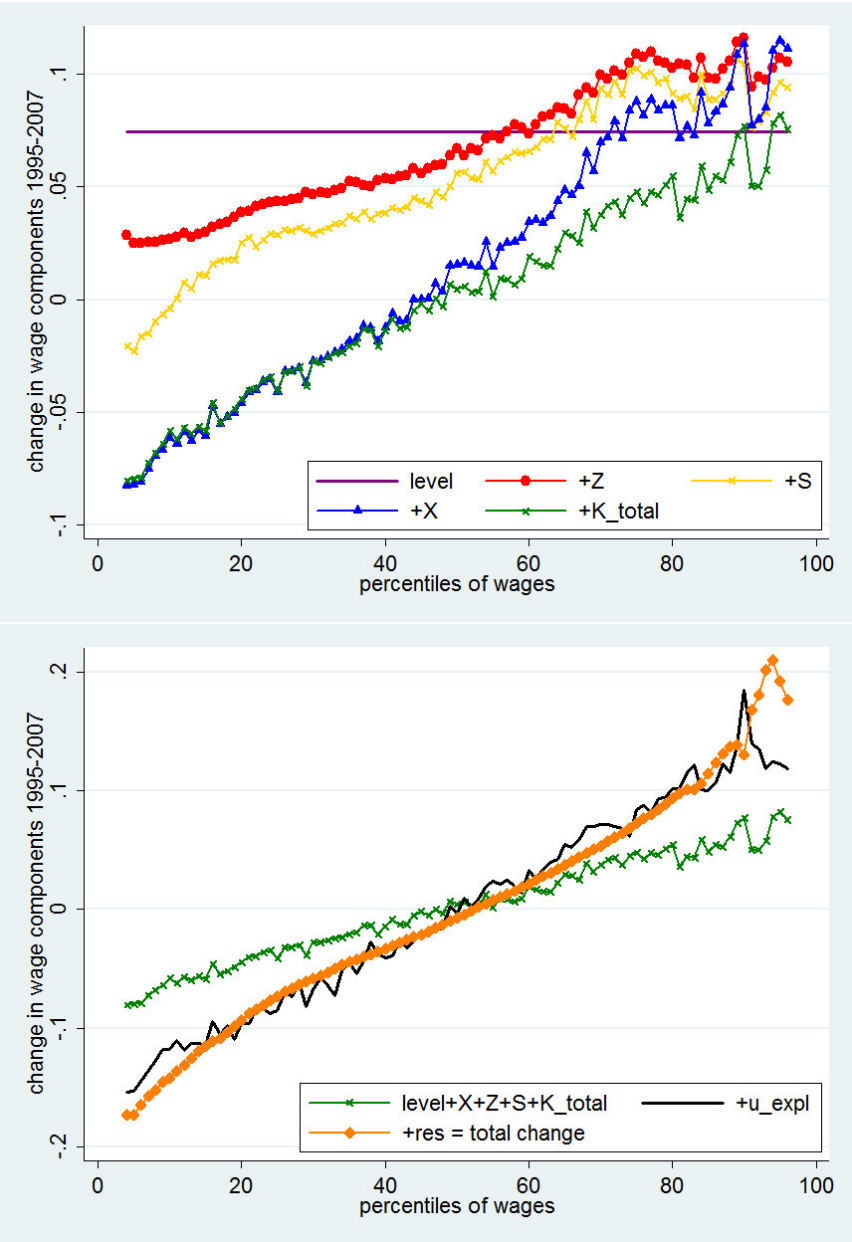
Notes: The figure shows the evolution of the 85/50 and 50/15 percentile ratio of the raw wage and the residual wage, each calculated with level values. The residual wage results after the subtraction of all estimated components of the wage regression, i.e., it corresponds to $\upsilon_{res} + \epsilon$, cf. the description of table 2.

Figure 4: (Adjusted) wage distributions



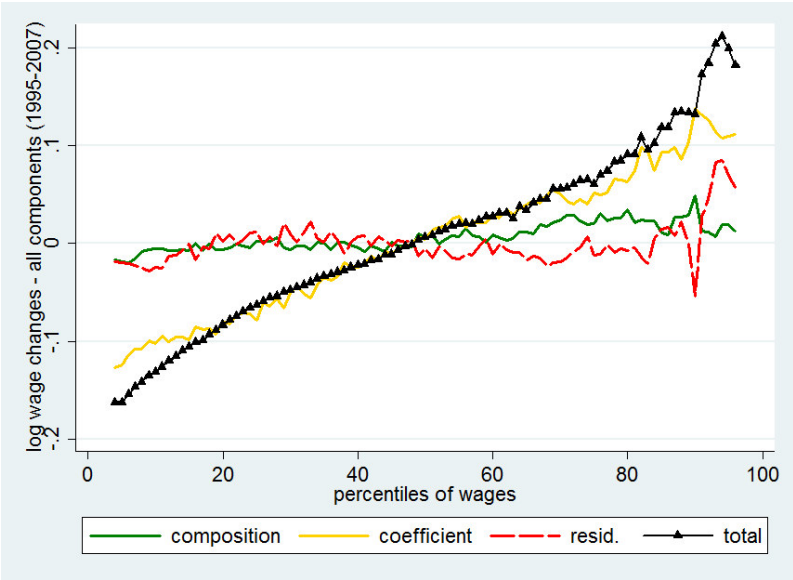
Notes: Both graphs show the distribution of daily wage levels in Germany in the years 1995 and 2007. In each of the graphs, the blue line represents the raw imputed wages as given in the data. The following hypothetical adjusted wages are constructed as follows. I simply subtract the effect of components from the log wage according to eq. (1) and re-transform this into a level value. Thus the red line is given by $exp(w_{it} - \hat{\beta}_x X_{it} - \hat{\beta}_w West_{kt} - \hat{\beta}_{xw} West_{kt} \cdot X_{it})$, and so on, as indicated in the legend. The yellow line appears, once all explainable components are subtracted.

Figure 5: Change in wage components 1995-2007



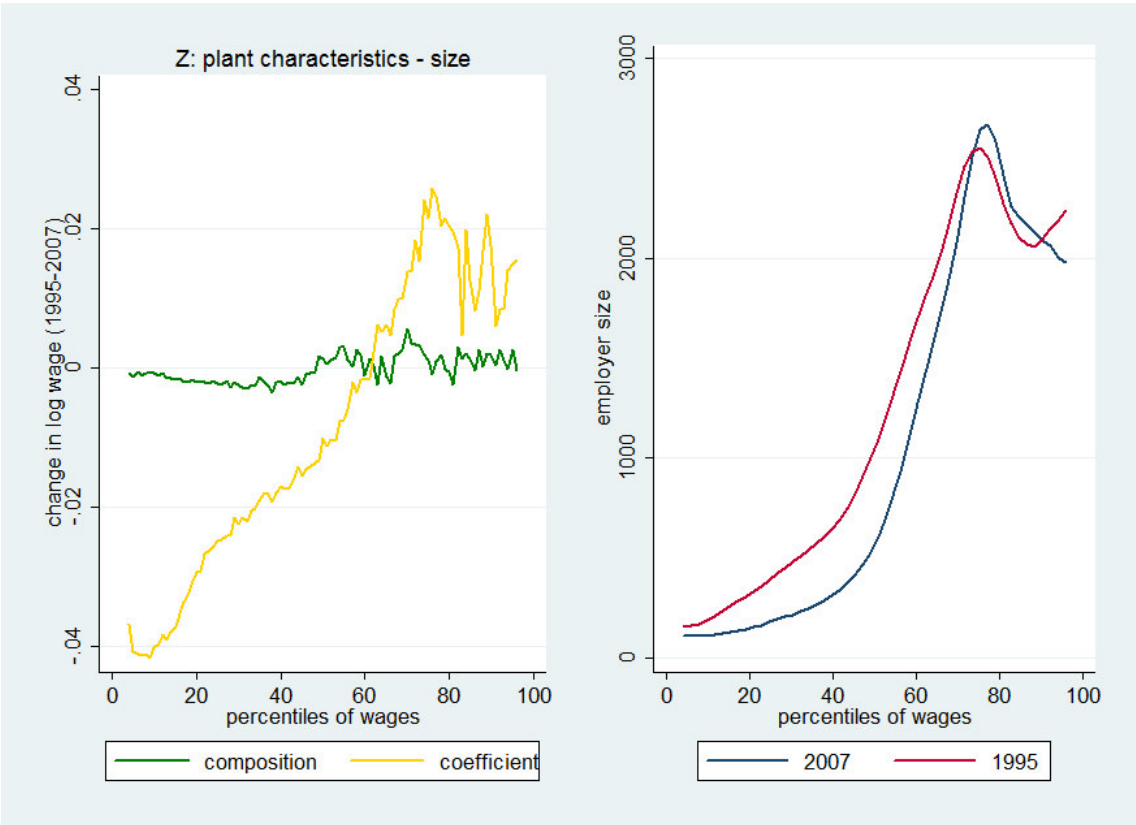
Notes: Both graphs show the distribution of composite changes in wage components between 1995 and 2007. Starting point is the change in levels ($\hat{\beta}_0 + \hat{\beta}_t D_t$), to which I subsequently add the change in the firm-specific component effect Z, the change in S, X and so on, as indicated in the graphs' legends.

Figure 6: Wage changes 1995-2007 - endowments vs. coefficients



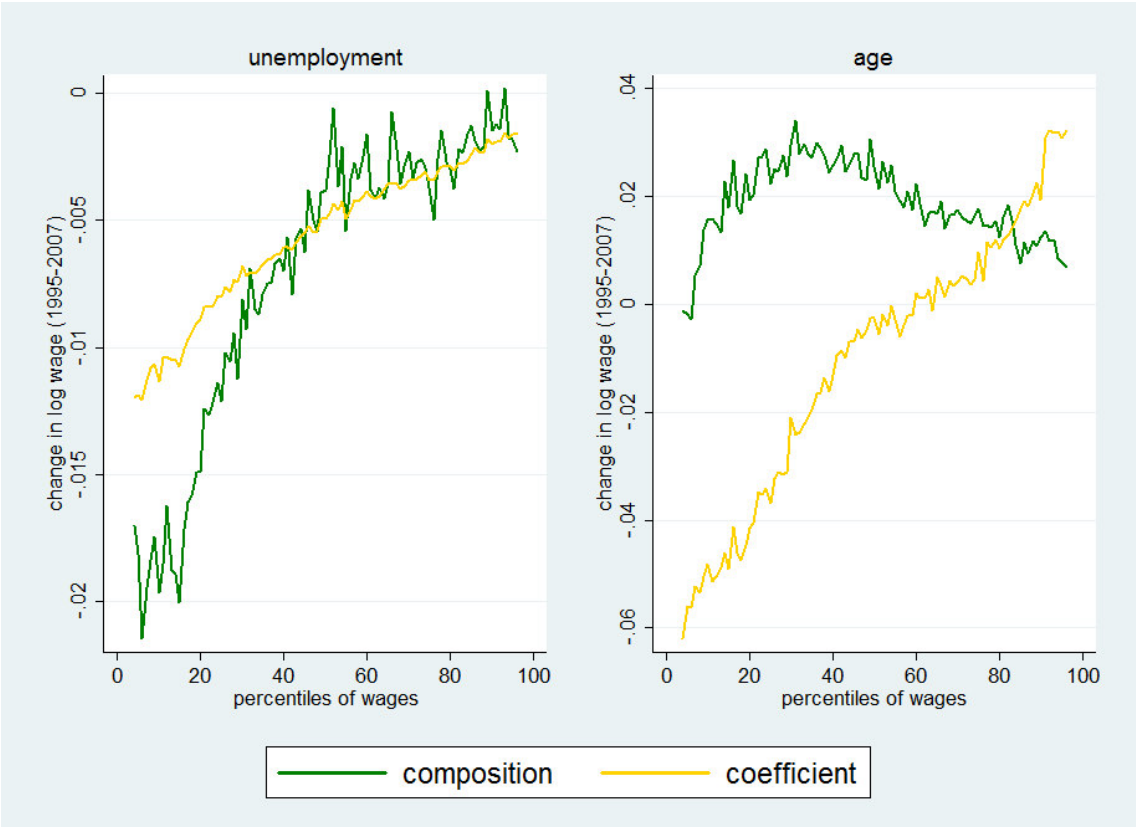
Notes: This figure shows how the total log wage change between 1995 and 2007 is separated into a endowment and a coefficient effect of all explanatory variables according to eq. (12). The change of the residual component is also displayed.

Figure 7: Endowment vs. coefficient changes - employer size



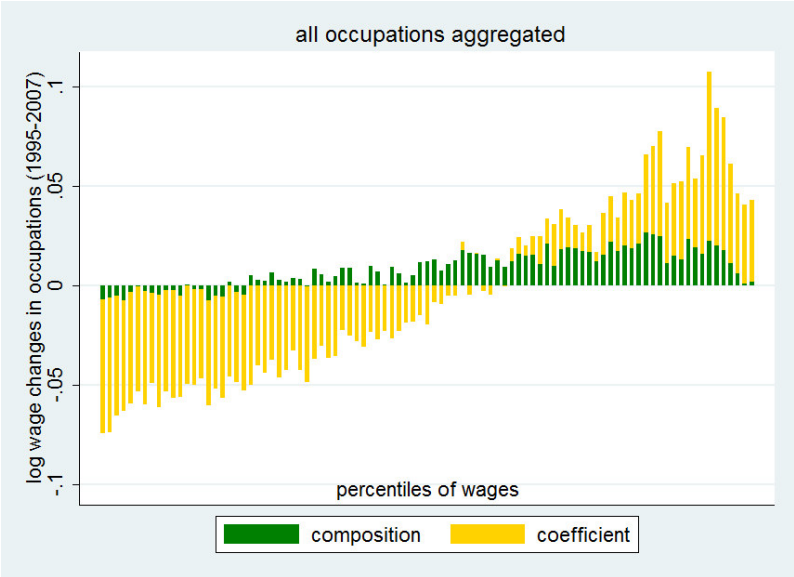
Notes: The left graph is analog to figure 6 but displays only the part of the aggregate change that is due to the employer size. The right graph shows the average employer size of workers along the wage distribution using a kernel-weighted local polynomial smoother.

Figure 8: Endowment vs. coefficient changes - unemployment and age



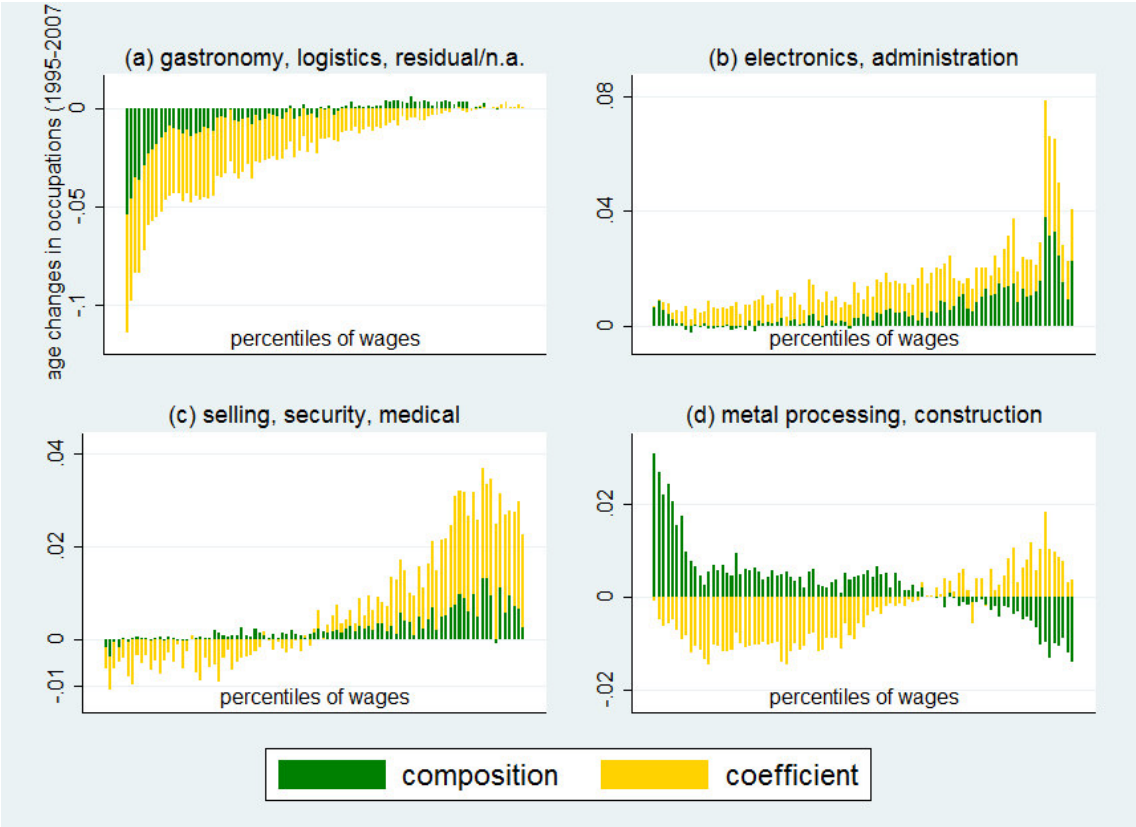
Notes: Both graphs are analog to figure 6. The left displays only the part of the aggregate change that is due to an unemployment spell prior to the current employment. The right graph shows the endowment and coefficient effects of the workers' age.

Figure 9: Endowment vs. coefficient changes - all occupations



Notes: The occupation-specific part of the wage regression is divided as in figure 6 and 7.

Figure 10: Endowment vs. coefficient changes - occupation segments



Notes: The 341 different occupations are classified into 21 segments according to Matthes *et al.* (2008). The four panels in the figure show the combined composition and coefficient effects in the 10 occupation segments with the largest changes over time.