Firm Size, Workforce Composition, and Wage Inequality

Daniel Bias[a], Chen Lin[b], Benjamin Lochner[c,d], Thomas Schmid[b]

[a] *Technical University Munich, Department of Financial Management and Capital Markets*
[b] *University of Hong Kong, Faculty of Business and Economics*
[c] *University of Erlangen-Nuremberg-Friedrich Alexander*
[d] *Institute for Employment Research (IAB)*

Abstract

Using a matched employee-establishment-firm data set covering German workers, we find that wage inequality increases monotonously with firm size: the largest firm decile pays 70% higher wages than the bottom decile, and their within-firm wage variance is 30% larger. Decomposing wage inequality into workforce composition and non-composition effects reveals that composition is responsible for three-quarters of the size-inequality relation within firms and half between firms. Higher wage variance in larger firms is largely explainable by more heterogeneous job characteristics and higher employee monitoring complexity. Higher wages in larger firms are not related to differences in profitability, monitoring complexity, or unionization levels, but different job characteristics and local labor markets play some role. Analyzing establishment size within firms reveals that larger establishments show more variation in workforce quality, but do not pay an economically significant wage premium.

☆We thank Christoph Kaserer, Dalia Marin, Christian Merkl, Stefan Seth, Sebastian Siegloch, Uta Schönberg, Elena Simintzi, and Stefanie Wolter for comments and suggestions.
Please send correspondence to Daniel Bias, E-Mail: daniel.bias@tum.de

August 26, 2019

1. Introduction

A recent study by the International Labour Organization found that the bottom half of workers were paid only 6.4% of global wages. This wage inequality, which rose sharply during the last decades (e.g., Katz and Autor, 1999), is nowadays one of the key political issues. In 2013, Barack Obama even described inequality as the "the defining issue of our time".[1] Only recently, the literature starts focusing on the role of firms, which ultimately set wages, for wage inequality in the economy (e.g., Mueller, Ouimet and Simintzi, 2017a; Bloom et al., 2018; Song et al., 2019).

In this paper, we analyze the role of firm size for wage inequality. Wage inequality can arise either because wages within larger firms are more dispersed ("within-firm inequality") or because larger firms pay their workers higher average wages ("between-firm inequality"). For our analysis, we use a linked employee-establishment-firm data set from Germany which covers about 50,000 individual firms, 120,000 establishments, and twelve million workers.[2] The uniqueness of this data set comes from the fact that it links administrative data about individual workers with firm-level information such as accounting figures.

Comparing the largest firm decile in terms of total assets to the bottom decile, we find that the largest firms have a 30% higher within-firm wage variance and pay 70% higher wages.[3] However,this positive effect of firm size is not only present when comparing the top and bottom of the size distribution, but increases monotonously with firm size. We use total assets as our main size proxy because it is most common in the financial economics literature and captures all resources available to the firm. However, for comparison, we also follow the majority of the inequality literature and use employee-based size proxies. Despite similar results for within-firm inequality, employee-based size measures tend to underestimate the effect of size on between-firm inequality.

[1] Washington Post, December 4, 2013.

[2] Worker-level information is based on administrative data which origins from the German social security system. The worker-level data and the employee-establishment-firm link is provided by the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB). Firm-level data is obtained from historical versions of Bureau van Dijk's Amadeus database.

[3] Firms in the largest decile have on average about 170m total assets and 500 employees, while those in the bottom decile have about EUR 0.5m total assets and 50 employees.

Next, we conduct a graphical investigation of the development of average firm size and wage inequality over time. We find that firm size increased substantially between 1995 and 2007 (i.e., the year before the financial crisis). This increase is not only observable for the largest firms, but also for smaller firms. For instance, the average size of both the largest 50 firms in Germany and firms ranked between 501 and 1000 in terms of size increased by about 70%. In the same time period, wage inequality approximately doubled, with a stronger increase for between than within inequality.[4] Thus, although this evidence is somewhat suggestive, it is consistent with the view that increases in average firm size contributed to the increasing wage inequality in the last decades.

To better understand why wage inequality is more pronounced in larger firms, we decompose between- and within-firm wage inequality into workforce composition and non-composition effects. Composition effects capture differences in workers' wages between small and large firms that exist due to differences in workforce quality. Non-composition effects are related to differences in wages for workers with the same quality. For the decomposition, we start by estimating an Abowd, Kramarz and Margolis (1999)-type regression (henceforth AKM) on the universe of full-time jobs held by workers between 20 to 60 years from 1993 to 2017.[5] Generally speaking, this approach exploits movers between establishments to disentangle the overall wage into different components (e.g., person and establishment fixed effects). We then use these AKM components to decompose the mean firm wage and the within-firm wage variance into workforce composition and non-composition effects. We find that composition effects play an important role for higher wage inequality in larger firms: the fact that larger firms employ, on average, more heterogeneous and higher quality workforce explains approximately three quarters of the within-firm inequality and half of the between-firm inequality.

Next, we try to explain why there is a positive relation between firm size and the different components of wage inequality. For composition effects, we hypothesize that larger firms have a higher workforce quality and heterogene-

[4]These patterns changes after the financial crises, with a decrease in firm size across the whole size distribution and a decreasing trend for wage inequality.

[5]Our approach follows Card, Heining and Kline (2013) (henceforth CHK) with some modifications (e.g., we also include female workers, see Sections 2.1 and 3.3 for more details).

ity because their job characteristics differ from smaller firms. It is often argued that larger firms have a greater capital intensity and that capital-skill complementaries exist (Hamermesh, 1980; Oi, 1983; Schmidt and Zimmermann, 1991; Lallemand, Plasman and Rycx, 2005). Thus, larger firms can make better use of employees' skills and are more likely to employ high-skilled workers. Furthermore, outsourcing of jobs that are not related to firms' core business and which often require workers with lower skills (such as food, cleaning, security, and logistics) became increasingly common over the past decades (Goldschmidt and Schmieder, 2017). If larger firms are more likely to engage in outsourcing, this reduces their reliance on low-skilled workers. Both effects can shift the average level of workforce quality upwards. Assuming that outsourcing is not completely possible and larger firms also need to rely on low and medium-skilled workers to support operations and that tasks in larger firms are more specialized, we also expect to see a greater heterogeneity of jobs in larger firms.

We find empirical support for this theoretical prediction: When we control for the fraction of (highly) complex jobs and the concentration of occupations within a firm, the impact of firm size on workforce quality dispersion becomes much weaker and nearly disappears. This indicates that the higher dispersion of workforce quality in larger firms is strongly related to differences in job characteristics between large and small firms. Furthermore, we find that measures for job complexity explain approximately one-quarter to one-third of the between-firm composition effect. Thus, differences in job characteristics explain at least partially the higher workforce quality in larger firms.

For the more dispersed wages after composition effects, we hypothesize that larger firms face more severe monitoring problems, which increases employees' opportunities for shirking. In order to reduce shirking behavior, larger firms can pay higher wages as incentive mechanism ("efficiency wage hypothesis").[6] Furthermore, more severe monitoring problems in larger firms may not only lead to higher average wages, but the wages could also be more dispersed if larger firms use more tournaments (i.e., wage differentials between hierarchy levels) and/or bonus payments (i.e., performance-based incentive payments) to incentivize workers.

[6] Theoretical motivation for this idea is, among others, provided by Eaton and White (1983) and Shapiro and Stiglitz (1984). Krueger and Summers (1988) are among the first to provide empirical evidence for the existence of noncompetitive wages.

Empirically, we find that the higher dispersion of wages in larger firms after workforce composition effects disappears when we control for monitoring complexity proxies (i.e., managers per employee and geographical dispersion of establishments within firms). Thus, more dispersed wages in larger firms seem to be strongly related to their higher monitoring complexity. However, we do not find empirical support that higher average wages in larger firms are related to their monitoring complexity.

Several explanations for higher average wages in larger firms after composition effects (the so-called "large-firm wage premium", LFWP) were discussed by prior literature.[7] As described in the previous paragraph, the efficiency wage explanation argues that larger firms face higher monitoring complexity and pay higher wages to incentive workers. Second, larger firms may generate higher rents by exploiting their market power, and they may share some of their higher rents with employees ("rent-sharing hypothesis").[8] Third, larger firms may have higher unionization rates (enabling employees to extract a higher fraction of rents) and a higher threat of unionization (Dickens and Lang, 1985). Fourth, larger firms may cluster in regions with high economic activity and thus a high demand for labor.

We fail to find any empirical support for the view that differences in monitoring complexity, profitability, or unionization-levels are the main drivers behind the existence of the LFWP. Neither controlling for monitoring complexity nor profitability nor industry-level unionization rates has a material impact on the relation between firm size and the average wage. However, when we compare small and large firms located in the same region which rely on the same labor market (i.e., those from the same industry), the LFWP is reduced by one-third or one-half, depending on the specification. Thus, it seems that factors related to local labor markets play some role for the existence of

[7]Possible explanations that are not covered in our discussion include differences between large versus small firm regarding their ability to screen workers' quality (Garen, 1985), working conditions (Brown and Medoff, 1989), governance structures (Pagano and Volpin, 2005; Cronqvist et al., 2009), ownership structures (Ellul, Pagano and Schivardi, 2018), internal labor markets (Tate and Yang, 2015), or other forms of employee participation like ESOPs (Kim and Ouimet, 2014).

[8]In this context, Christofides and Oswald (1992) find that real wage is positively related to past industry performance, which is consistent with the rent-sharing hypothesis. Related to this finding, Abowd and Lemieux (1993) use product market competition to show that rent-sharing considerations affect workers' wages. Card et al. (2018) provide a comprehensive discussion of the literature linking firm productivity to wages.

the LFWP.

To summarize, our tests indicate that more heterogeneous job characteristics in larger firms and higher employee monitoring complexity explain, to a high degree, their higher within-firm wage inequality. Between-firm inequality, i.e., higher average wages in larger firms, are more challenging to explain. Differences in job characteristics can partly explain larger firms' higher workforce quality, but a significant part of the higher average workforce quality in larger firms remains even after controlling for this factor. Local labor market factors seem to explain some of the higher average wages in larger firms after composition effects, but a substantial part remains unexplained. While we acknowledge that our proxies (especially for monitoring) are far from being perfect, we find that efficiency wages, rent sharing, and unionization are unlikely to be the main driver behind this LFWP.

To better understand the role of firm size for wage inequality, we also investigate how the size of establishments within firms affects wages and their dispersion. When we compare establishments of the same firm, we find that larger establishments pay more dispersed wages. This result is mainly driven by more heterogeneous workforce quality in larger establishments. Similar to our firm-level results, it thus seems that establishment size is positively related to the heterogeneity of jobs. Although there is some evidence that larger establishments employ higher quality workforce, the economic significance of this result is low compared to differences between large and small firms. After controlling for composition effects, larger establishments within the same firm pay only marginally higher wages than smaller establishments: compared to differences between firms, the size effect within firms is reduced by 90%. Thus, it seems that larger firms pay higher average wages, relatively independently of the size of a particular establishment within that firm. By contrast, the size of an establishment within a (large) firm plays a very significant role for the dispersion of wages, and especially for the dispersion of workforce quality.

Our results contribute to several strands of the literature. Most importantly, we add to the literature which investigates how firms affect wage inequality. Despite a recent literature highlighting that firms matter for increasing wage inequality[9], it is still not well understood how and especially why

[9]See, among others, Alvarez et al. (2018) for Brazil, Card, Heining and Kline (2013) for Germany, Card et al. (2018) for Portugal, Faggio, Salvanes and van Reenen (2010) for the

firm size affects wage inequality.

With regard to higher dispersion of wages in larger firms, Davis et al. (1991) were among the first to document a positive correlation between plant size and wage dispersion. Mueller, Ouimet and Simintzi (2017a) and Mueller, Ouimet and Simintzi (2017b) find that the overall within-firm wage inequality is more pronounced in larger and more profitable firms, but they do not decompose wage inequality in workforce composition and non-composition effects. We complement their analysis by showing that composition effects account for three-quarters of the within-firm pay inequality. Song et al. (2019) investigate the contribution of firms to the rise in wage inequality. Although they do not focus on the role of firm size, they show that the rise of within-firm wage inequality happened mainly in mega firms with more than 10,000 employees. Our paper complements their analysis by showing that wage inequality increases monotonously with firms size (not only in mega firms) and by providing (some) explanations for the positive relation between firm size and wage inequality.

For between-firm wage inequality, there is a large literature which documents higher average wages in larger firms (e.g., Brown and Medoff, 1989; Oi and Idson, 1999; Bloom et al., 2018). However, this literature often lacks firm-level data such as profitability or establishment structures, which makes the testing of several potential explanations for higher wages in larger firms (e.g., rent-sharing) challenging. Our contribution to this literature is to test whether such firm-level factors help explaining the LFWP. Somewhat surprisingly, our results do not show any evidence that rent sharing, monitoring complexity, or unionization have a substantial impact on the LFWP.

Our results also contribute to the inequality literature by documenting the following stylzed facts. First, average firm size increased between 1995 and the financial crisis not only for the largest firms, but across the whole size distribution. This is in line with the recently documented rise of superstar firms (Autor et al., 2019) and the finding that large public U.S. firms tripled their average size in the last two decades (Grullon, Larkin and Michaely, 2019). Second, our firm-level data allows us to use firm capital as size measure.[10] Capital-based and employee-based firm size measures lead to relatively sim-

U.K., and Song et al. (2019) for the U.S.

[10]The fall of the labor share in the largest firms (Autor et al., 2019) highlights the importance to differentiate between firm size based on capital and labor.

ilar results for within-firm wage inequality, but employee-based proxies seem to underestimate the effect of size on between-firm wage inequality. Third, between-firm wage differentials are mainly driven by firm size, not establishment size. On the other hand, within-firm wage inequality is mainly driven by establishment size, not firm size. These findings show that it is important to distinguish between establishment and firm size when analyzing the relation between size and wage inequality.

2. Data

2.1. Employer-employee Data

We use administrative linked employer-employee data provided by the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB) of the German Federal Employment Agency (Bundesagentur für Arbeit). The Integrated Employment Biographies (IEB) data originates from records from the German social security system. The data includes total earnings and days worked at each job in a year, as well as information on education, occupation, industry, and part-time or full-time status.[11]

The data preparation follows the steps conducted by CHK. We obtain information on the universe of full-time jobs held by workers with age 20-60 from 2010 to 2017.[12] While CHK examine male workers in West Germany, we cover both male and female workers in East and West Germany. We exclude marginal employment and apprenticeship. As in CHK, we focus on the main job held by each worker in a given year, that is the job with the highest total wage sum (including bonus payments). For all of these jobs, we calculate the average daily wage by dividing the total wage sum by the total duration of the main job. One important limitation is that wages are only reported up to a time and region specific threshold – the contribution assessment ceiling ("Beitragsbemessungsgrenze"). We follow the procedure suggested by Dustmann, Ludsteck and Schönberg (2009) and CHK and impute the upper tail of the wage distribution by running a series of Tobit regressions, allowing for a maximum degree of heterogeneity by fitting the model separately for gen-

[11] For further details on the data set, please refer to the technical report by Antoni, Ganzer and vom Berge (2016).

[12] The administrative data origins from the social security system, so the IEB data does not include employment spells of civil servants and self-employed workers.

der, time, education levels, and eight five-year age groups. We impute missing and inconsistent information in the education variable using the methodology proposed in Fitzenberger, Osikominu and Völter (2006).

The employee-establishment data set consists of over 165 million worker-establishment-year observations, almost 32 million unique workers, and more than 2.7 million establishments.

2.2. Firm-establishment Data

The linked employer-employee data provides only information on employees and establishments. E.g., individual i is employed at establishment j. To add information on the firm structure, we use the novel ORBIS-ADIAB data set that links the establishments identifiers of the Institute for Employment Research to the Bureau van Dijk (BvD) identifier as firm-level identifier. E.g., establishment identifiers j_1 and j_2 belong to the firm k. Firm-level financial information from 2010 to 2016 is then derived from the Amadeus database by BvD.

This linkage process of establishments to firms is described in full detail by Antoni et al. (2018). To identify all matches between establishments which truly belong to same company, they use various variables of which the establishment and the company name as well as the legal form are the most important ones. After extensive pre-processing, they apply record linkage techniques. One limitation of this data set is that the company information for the record linkage is extracted on January 30, 2014. Since our main analysis is based on the 2010-2017 period, we assume this limitation to have a minor effect on our results.

2.3. Sample Construction

We restrict the sample to firms with more than 20 employees in the IAB data to ensure that statistics about wage variations within firms are meaningful. We also exclude firms from our sample for which we do not obtain information on total assets and sales.[13] The final sample covers a total of

[13]Further, we exclude establishments for which we cannot estimate the AKM components. The AKM approach identifies establishment fixed effects by workers moving between establishments and are estimated relative to an omitted reference establishment. The estimation is performed on the largest set of establishments that are connected through worker transitions.

49,069 firms with 119,641 establishments and about 12.4 million individual workers.

2.4. Descriptive Statistics

Table 1 provides descriptive statistics on the firm level. Unscaled financial variables are consumer price index adjusted and all continuous variables are winsorized at the 1st and 99th percentiles. We report firm mean values in the 2010-2017 period. The average firm employs about 67 workers in Germany and about 84 employees worldwide. It has mean total assets of EUR 5.7m, a sales of EUR 233,000 per employee, an operating profit of EUR 22,000 per employee, and a return on assets of 13%.

3. Method

3.1. Between and Within Wage Inequality

$y_t^{i,j,k}$ denotes the log of the real daily wage of worker i employed by establishment j belonging to firm k in year t.[14]. Using information on the firm structure, we construct,

$$\begin{aligned} \bar{y}^k &= \frac{1}{N_k} \sum_t \sum_i y_t^{i,j,k} \\ var_k(y_t^{i,j,k}) &= \frac{1}{N_k} \sum_t \sum_i (y^{i,k} - \bar{y}^k)^2 \end{aligned} \tag{1}$$

where $\bar{y}^k$ is the mean wage of firm k, $var_k(y_t^{i,j,k})$ the variance of wages within firm k, and N_k the total number of employee-year observations of firm k in our sample. In parts of our analysis, we run analysis on establishment level. Then, we construct,

$$\begin{aligned} \bar{y}^j &= \frac{1}{N_j} \sum_t \sum_i y_t^{i,j} \\ var_k(y_t^{i,j}) &= \frac{1}{N_j} \sum_t \sum_i (y^{i,j} - \bar{y}^j)^2 \end{aligned} \tag{2}$$

[14]For notational convenience, we suppress the dependence of subscript j on worker i and t, such that j = J(i,t), as well as the dependence of subscript k on worker i, establishment j and t, such that k = K(i,j,t).

where $\bar{y}^j$ is the mean wage of establishment j, $var_k(y_t^{i,j})$ the variance of wages within establishment j, and N_j the total number of employee-year observations of establishment j in our sample.

3.2. Wage Inequality and Size

We are interested in the greater variation of wages within larger firms compared to smaller firms ("within-firm inequality") and the wage differences between smaller and larger firms ("between-firm inequality"). To establish these relations, we regress, in our baseline models, the firm mean wage and the within-firm variance of wages on firm size,

$$\begin{aligned} var_k(y_t^{i,j,k}) &= \beta size_k + \epsilon_k \\ \bar{y}^k &= \beta size_k + \epsilon_k \end{aligned} \tag{3}$$

where all models are weighted by firms' total number of (German) employees. We cluster standard errors on firm level. Our baseline measure of firm size is total assets. In the Appendix C, we show the robustness of our main results to the use of alternative employee-based measures, total number of German employees and worldwide employees.

To analyze the role of establishment size within firms, we estimate regression models using establishment-level data, e.g.,

$$var_k(y_t^{i,j}) = \beta size_j + \pi_k + \epsilon_j \tag{4}$$

where π_k denotes the firm fixed effect to compare only establishments within one firm. We measure establishment size by the establishments' number of employees. All models on establishment-level are weighted by the establishments' number of employees. Standard errors are still clustered on firm level.

3.3. Implementation of AKM-type Regression Model

We follow the CHK implementation of the model introduced by AKM. We estimate the following regression model in the 2010-2017 period,

$$y_t^{i,j} = \alpha^i + \psi^j + \beta X_t^i + r_t^{i,j}, \tag{5}$$

where $y_t^{i,j}$ denotes the log average daily wage of individual i in year t, α^i the person fixed effect, ψ^j an establishment fixed effect, X_t^i an index of time-varying observable characteristics and $r_t^{i,j}$ an residual. X_t^i includes an unrestricted set of year dummies as well as quadratic and cubic terms in age fully interacted with educational attainment.[15]

The estimation is done on the largest connected set of establishments that are linked by worker transitions within the 2010-2017 period. Importantly, we carry out the estimation on the universe of full-time linked employer-employee data set for Germany and not only on the subset for which we also observe the firm structure. Our data preparation follows the steps conducted by CHK. The most relevant modifications are that we also cover workers in East Germany and female workers (for further details, see Section 2.1).

3.4. Decomposition of Firm Mean Wage and Within-Firm Variance of Wages

We use the parameter estimates from the AKM-type regression to decompose the firm mean wage ($\bar{y}^k$) and within-firm variance of wages ($var_k(y_t^{i,j,k})$). Ignoring time-varying working characteristics βX_t^i for now, the firm-level mean can be decomposed into,

$$\bar{y}^k = \bar{\alpha}^k + \bar{\psi}^k \tag{6}$$

where means are calculated over all person-year observations of firm k in our sample. $\bar{\alpha}^k$ is the composition part, measuring the average worker quality, identified from time-invariant worker characteristics (the worker fixed effects). $\bar{\psi}^k$, the firm-level mean of the establishment fixed effects, to which we refer as "firm fixed effect". It reflects firm-specific pay policies like rent sharing (e.g., Card et al., 2018) or efficiency wages (e.g., Krueger and Summers, 1988).

Song et al. (2019) extend the standard AKM / CHK variance decomposition to differentiate between a within-firm component and a between-firm component. We adopt the within-firm component to our setting, that considers the firm-establishment structure. Again ignoring time-varying working characteristics βX_t^i for now, we can decompose the within-firm variance into,

$$\begin{aligned} var_k(y_t^{i,j,k}) &= var_k(\alpha^i - \bar{\alpha}^k) + var_k(\psi^j - \bar{\psi}^k) \\ &\quad + 2cov_k(\alpha^i - \bar{\alpha}^k, \psi^j - \bar{\psi}^k) + var_k(r_t^{i,j}) \end{aligned} \tag{7}$$

where the variances are calculated over all person-year observations of firm

[15]As in CHK, we normalize the age variable to 40 years. See Card et al. (2018) for a discussion.

k in our sample. The first component is attributable to within-firm differences in worker quality. The second part measures differences in establishment-specific pay policies like an establishment-size dependent wage premium (e.g., Brown and Medoff, 1989). The third part reflects sorting of workers within the firm, e.g., high paid workers might be sorting to high-paying establishments of a firm. The last term is the idiosyncratic wage component which is orthogonal to all other components (Card, Heining and Kline, 2013).

4. Results

4.1. Firm Size and Wage Inequality

As a first step, we investigate how firm size[16] affects wage inequality within and between firms. We show this in a graphical and a regression analysis.

Figure 1 illustrates the relation between wage inequality and firm size. For this figure, we sort all sample firms into deciles according to their total assets. Each decile consists of about 4,900 firms. Subfigure (a) focuses on within-firm wage inequality and presents the mean variance of wages within firms in each decile. Additionally, we also show the mean within-firm variances of the different AKM components. However, for simplicity, we only show the within-firm variances of the AKM components and ignore their covariances. Hence, this figure does not show a full decomposition of the within-firm variance (see Appendix D for a full decomposition).

We find a positive correlation between firm size and the within-firm variance. For the bottom decile of firms, the wage variance is on average 0.090, while it is 0.115 for the top decile. In relative terms, this equals an increase of 28%.[17] The findings that the variance of the person fixed effects accounts for about 80% of the wage variance and that both move nearly parallel suggest that the greater variation of wages within larger firms is mainly related to their higher heterogeneity of workforce quality. The second most important component of overall wage variance is the variance of the residual. The variance of the residual includes all aspects that are not captured by the AKM model and

[16] As explained before, we use the total assets of a firm as our main size proxy. In Section 4.2, we will compare this size measure to employee-based proxies.

[17] Using the difference between the 90th percentile and the 10th percentile as alternative measures for wage dispersion within firms, we find that the relative increase between the bottom and top decile is 25% (results unreported).

can be interpreted as the variance of idiosyncratic worker premiums. However, the graphical analysis fails to provide evidence that the variance of the residual increases with firm size. Rather, the relation between firm size and the variance of the residual seems to follow a U-shape. Thus, factors that are not captured by the other AKM components do not seem to drive the positive relation between firm size and within-firm wage inequality.

— Figure 1 about here —

Subfigure (b) focuses on between-firm wage inequality and presents the average wage paid by a firm for each size decile. Additionally, we again show the decomposition of the overall wage into the AKM components. In line with the prior literature, we find that the average wage is increasing substantially over the size groups. While firms in the bottom size decile pay wages that are 35% below the sample mean, firms in the top decile pay wages that are 35% above the sample mean. Thus, the difference in terms of average wages between firms in the bottom and top decile is approx. 70%. The decomposition shows that about 40% of this differences is due to workforce composition effects and the remaining 30% due to firm-specific wage premiums (i.e., the premium paid to a hypothetical worker with average observable and unobservable characteristics).

Table 2 presents the results of the corresponding regression analysis. In Panel A, we use the within-firm variance of wages and AKM components as dependent variables. The independent variable in all regressions is the log of a firm's total assets (see Equation 3). For the within-firm variance of wages, the regression coefficient is 0.0045. This implies that a firm would increase its within-firm inequality by 4.1% in relative terms if its size doubles.[18] Analyzing the AKM components reveals that the person fixed effects, i.e., factors related to workforce composition, account for three-quarters of the positive relation between size and wage variance within firms. Approximately one-fifth of the wage variance is explained by the variance of the residual, i.e., variance of the idiosyncratic worker premiums that are not related to workforce composition or other components in the AKM.[19]

[18]Note that employee-weighted mean of the within-firm variance is 0.110.

[19]Please note that the within-firm variance of log wages can be decomposed into within-firm variances of the AKM components and their covariances. The regression coefficients of the within-firm variances of the AKM components do not add up to the regression coeffi-

— Table 2 about here —

In Panel B, we focus on wage inequality between firms and use the firm mean wage and its AKM components as dependent variables. The independent variable in all regressions is the log of a firm's total assets (see Equation 3). For the firm mean wage, the regression coefficient is 0.12. This implies that a firm would increase its average (log) wage by 12% in relative terms if its size doubles. The coefficients for the person fixed effect is 0.067. Thus, slightly more than half of the positive relation between firm size and average wages can be explained by workforce composition effects, i.e., higher workforce quality in larger firms. The coefficient for the firm fixed effect, which is nothing else than the average across all establishment fixed effects of a firm, is 0.051. Thus, firm-specific wage premiums account for slightly less than half of the higher wage in larger firms.[20]

4.2. The Measurement of Firm Size

Our main measure of firm size is total assets. The majority of the previous literature in the context of wage inequality used employees to measure (firm) size. This is at least partly related to the fact that accounting data, such as total assets, is normally not available in dataset that cover individual workers. Our data set includes firm-level accounting data from Amadeus which allows us to use total assets as our main proxy for firm size. We use this measure because it is the most commonly used size proxy [21] and it captures all resources available to the firm. Thus, total assets is a more commonly used and general firm size measure than the number of employees.

Nevertheless, we also use the number of a firm's worldwide employees according to the Amadeus database or the number of employees in Germany according to the data provided by the IAB as size proxies. First, we plot the

cient of the within-firm variance of wages in our table because the covariances of the AKM components are not reported.

[20]For the residual, the coefficient on total assets is omitted or virtually zero. Therefore, we report it as not relevant (n.r.). Please note that we only present the first two true digits of coefficient estimates. Nevertheless, the the regression coefficients of the AKM components without rounding add up to the regression coefficient of the firm mean wage.

[21]Dang, (Frank) Li and Yang (2018) analyze the 100 most cited papers from top economics, finance, and accounting journals that use firm size measures in the area of empirical corporate finance. Among those 100 papers, 87 use a single size measure, which is total assets in 49 papers, market capitalization in 20 papers, sales in 16 papers, and number of employees in 2 papers.

relation between wage firm size and wage inequality in Figure 2 and use total assets or the employee-based measures to sort firms into size deciles. Subfigure (a) shows that the employee-based size measures lead to very similar results as total assets for within-firm wage inequality. For between-firm wage inequality in Subfigure (b), we find that the inequality-size relation is substantially weaker when using employee-based size measures. While there is a 70% difference between firms in the top and bottom size decile according to their total assets, this difference is only about 30% when we sort firms by their number of employees.

We also repeat our regressions from Table 2 using the employee-based size measures. The results, which are reported in Appendix C, are not substantially different from our main regression results. However, we again find that the economic magnitude in the between-firm regressions is smaller when using employee-based size measures. Overall, these results indicate that the choice of the size proxy does matter in the context of wage inequality and that employee-based proxies can under-estimate the impact of firm size on between-firm wage inequality.

— Figure 2 about here —

4.3. The Development of Firm Size and Wage Inequality over Time

To shed some light on the question whether increases in firm size over the last decades can help to explain the rise in wage inequality, we graphically analyze the development of firm size and wage inequality over time. This test uses an extended sample period from 1995 to 2015.[22] For firm size, we do not only focus on the largest firms, but also investigate how the size of smaller and more typical firms changed over time. Thus, we distinguish between firms ranked 1 to 50, 51 to 100, 101 to 250, 251 to 500, and 501 to 1000 in terms of firm size in each year.

Subfigre (a) of Figure 3 shows how the average size of firms in these bins developed over time. We find that firm size increased substantially between 1995 and the financial crisis. This increase is not only observable for the largest firms, but also for smaller firms. For instance, the average size of both the largest 50 firms in Germany and firms ranked between 501 and 1000 in terms

[22] In our other tests, we focus on the years 2010 to 2017 (see Sections 2.2 and 2.1 for a detailed explanation).

of size increased by about 70%. Between 2007 and 2010, firm size declined for all bins, and it remained relatively constant after 2010. The development of within-firms, between-firms, and total wage inequality is shown in Subfigure (b). Similar to the development of firm size, we find an upward trend of wage inequality between 1995 and 2008. The overall wage inequality approximately doubled in this time period, which was mostly driven by increasing between-firm inequality. Between 2009 and 2010, wage inequality is relatively constant, and it shows a declining trend between 2011 and 2015. One possible explanation of the sharp decline between 2014 and 2015 is the introduction of the mandatory minimum wage in Germany on January 1st, 2015.

When comparing the development of average firm size and wage inequality, it is remarkable that both increased in a relatively similar manner until the financial crisis. Although both stopped their upward trend since the financial crisis, their development was somewhat different with a stagnation of firm size and a reduction of wage inequality. Although this evidence is somewhat suggestive, it is consistent with the view that increases in average firm size contributed to the increasing wage inequality in the last decades.

4.4. What Explains Higher Wage Inequality in Larger Firms?

In this section, we test possible reasons for the positive relation between firm size and wage inequality. In the Section 4.1, we have established that higher within-firm wage inequality in larger firms is driven by the larger heterogeneity of their workforce quality and—to a smaller extent—higher variance of the idiosyncratic worker premiums. The between-firm inequality stems from higher average workforce quality in larger firms and firm-specific wage premiums. We explain each of these four components in a separate subsection below.

4.4.1. Within-firm Inequality: Heterogeneity of Worker Quality

Our main hypothesis is that larger firms have a higher heterogeneity of workforce quality because their job characteristics differ from those in smaller firms. Because larger firms can make better use of workers' skills due to their higher capital intensity (Hamermesh, 1980; Oi, 1983; Schmidt and Zimmermann, 1991; Lallemand, Plasman and Rycx, 2005), we expect that occupations which require (very) high skills are clustered in these firms. At the same time, larger firms still need to rely on occupations that require medium or low skilled workers to support their operations. Second, despite recent outsourcing trends (Goldschmidt and Schmieder, 2017), larger firms are more likely to perform tasks which are not their core business in-house due to economies of scale (e.g., tax reporting). Third, occupations in larger firms are more focused on a specific task because the firm can divide the tasks into smaller parts, creating higher levels of specialization (Bryson, Erhel and Salibekyan, 2017). All these factors likely increase the heterogeneity of occupations in larger firms and—as a consequence—the heterogeneity of workforce quality.

Empirically, we use three different proxies to measure the dispersion of occuptations within firms. The first is the number of employees who perform complex tasks standardized by total employees.[23] The second measure is closely related to the first and captures the fraction of highly complex occupations. The last proxy uses the number of different occupations within a firm to measure the concentration of occupations. For this purpose, we calculate the Herfindahl index on firm-level based on the share of occupations within the firm.[24] We plot how these three proxies—and the number of occupations—are related to firm size in Figure B.2. The fraction of (highly) complex occupations and the number of occupations increases with firm size. For example, firms in the top size decile have on average 40 different occupations, while the corresponding number for firms in the bottom decile is only 15. Not surprisingly, the concentrations of occupations as measured by the Herfindahl index falls with firm size. Thus, job characteristics seem to differ substantially between small and large firms, potentially explaining the higher variation of workforce quality in the latter.

As a next step, we include these three proxies as control variables in our baseline regressions. Table 3 presents the results. For comparison, Column 1 shows the baseline model, regressing the within-firm variance of person fixed effects on total assets. The coefficient estimate is 0.0034. In Column 2, we add the share of employees carrying out complex task to the model. This reduces the coefficient estimate on firm size to 0.0026. In Column 3, we include the share of employees performing highly complex tasks. The coefficient estimate

[23]The Classification of Occupations 2010 (KldB 2010) makes it possible to differentiate between occupational activities according to four degrees of complexity, which are unskilled or semi-skilled, specialist, complex specialist, and highly complex activities. The level of requirement is given in the last digit of the 5-digit occupation code.

[24]For this calculation, an occupation is measured by the first 3-digits of the Classification of Occupations 1988 (KldB 1988).

on firm size drops to 0.0018. In Column 4, we add the Herfindahl index of occupations to the model. The coefficient on firm size is 0.0023 in this model, which is also considerably smaller than in the baseline model. The effects of the additional variables are as expected: a higher share of complex occupations increases the heterogeneity of the workforce quality, while the opposite is true for a higher concentration of occupations. The last column presents a specification in which we include all three proxies simultaneously. The coefficient on firm size is now 0.00075, which is 78% smaller than in our baseline model. The t-value drops from a highly statistically significant value of about eight to a marginally significant value of less than two. Thus, differences in job characteristics seem to be a key driver for the greater worker heterogeneity in larger firms.

—Table 3 about here —

4.4.2. Within-firm Inequality: Variance of Idiosyncratic Worker Premiums

Idiosyncratic worker premiums are wages after accounting for all AKM factors, including workforce quality. We hypothesize that larger firms exhibit a higher variance of these premiums due to more severe monitoring problems. Because monitoring is more complex in larger firms, employees have more opportunities for shirking (e.g., Eaton and White, 1983; Shapiro and Stiglitz, 1984). In order to reduce shirking behavior, larger firms can use more tournaments (i.e., wage differentials between hierarchy levels) and/or bonus payments (i.e., performance-based incentive payments) to incentivize workers. Thus, wages can be more dispersed in larger firms, even after accounting for their higher heterogeneity of workforce quality.

To measure monitoring complexity, we use three different proxies. Our first measure is managers per employees, which captures the relation between (supervising) managers and normal workers. We assume that a high fraction of managers, relative to workers, indicates higher monitoring complexity. The other two measures we use exploit the geography of a firm's establishments. Monitoring is likely more complex if the median distance between establishments and the headquarter is higher and if their location is more dispersed (e.g., Giroud, 2013). Thus, we use the median distance and the standard deviation of the distances as second and third measures. We plot how these three proxies are related to firm size in Figure B.3. All three show a positive relation with firm size. For example, the fraction of managers is about two percent in the bottom firm size decile, while it is ten percent in the top decile. Although we acknowledge that these proxies are far from being perfect, they indicate that monitoring complexity increases with firm size.

Next, we include these three proxies as control variables in our baseline regressions (see Table 4). The dependent variable is the within-firm variance of idiosyncratic worker premiums, but we excluding managers for its calculating.[25] For comparison, Column 1 shows the baseline model, regressing the variance of idiosyncratic worker premiums on total assets. The coefficient estimate is 0.00077. In Column 2, we include the ratio of managers to employers as control variable. In this specification, the coefficient on total assets drops by half to 0.00038. In Column 3, we control for the median distance between the firm headquarter and the establishments. Here, the coefficient estimate drops even further to 0.0025. Column 4 add the standard deviation of this distance to the model. The coefficient of total assets is slightly lower than in the baseline model (0.00064). In the last column, we control for all three proxies of monitoring complexity at the same time. In this specification, the coefficient estimate on size is negative (-0.000061) and statistically insignificant. The coefficient estimates for our monitoring complexity proxies are as expected as all three have a positive impact on the variance of idiosyncratic worker premiums. Overall, this test provides evidence that the greater variation of idiosyncratic worker premiums in larger firms is related to their higher monitoring complexity.

—Table 4 about here —

4.4.3. Between-firm Inequality: Worker Quality

We hypothesize that larger firms have a higher (average) workforce quality because their job characteristics differ from those in smaller firms. The argument here is similar as for their higher heterogeneity of workforce quality. First, capital-skill complementarities lead to a higher share of occupations which are (highly) complex and require workers with high skill levels. Second, larger firms may be more likely to engage in outsourcing, which reduces

[25]Managers are excluded to mitigate the concern that the higher share of managers itself and not the higher monitoring complexity in larger firms is driving the greater variation of the AKM residuals. The results are very similar if we do not exclude managers.

occupations which require lower skills. Examples include food, cleaning, security, and logistics tasks (see Goldschmidt and Schmieder, 2017, for more details on outscoring trends in Germany). Both factors can lead to a higher share of "complex" occupations in larger firms, which shifts the average level of workforce quality upwards.

We use same proxies as in Section 4.4.1 to measure the complexity of occupations. These are the fraction of employees conducting complex and highly complex tasks and the dispersion of occupations within firms. As expected, larger firms have a higher fraction of (highly) complex occupations and a lower concentration of occupations (see Figure B.2).

In Table 5, we include these variables as controls in our baseline regressions. For comparison, Column 1 shows the baseline model, regressing the mean worker quality, measured by the average person fixed in a firm, on total assets. The coefficient in the baseline model is 0.067. In Column 2, we add the share of employees carrying out a complex task to the model. The coefficient on total assets decreases to 0.056. Next, we include the share of employees carrying out highly complex tasks. This shrinks the coefficient on total assets even further to 0.052. In Column 3, we add the Herfindahl index to the model. This reduces the coefficient on total assets marginally to 0.062. The effects of our measures for job complexity are as expected, with a higher average worker quality in firms that have a higher share of (highly) complex occupations. Column 5 presents a joint model which includes all three control variables simultaneously. The coefficient of total assets is 0.044 in this model, which represents approximately a one-third reduction compared to the baseline model. Hence, differences in the complexity of occupations can partially explain the higher workforce quality in larger firms.

—Table 5 about here —

4.4.4. Between-firm Inequality: Large Firm Wage Premium

Previously we documented that firm-specific wage premiums account for slightly less than half of the higher wage in larger firm. This means that larger firms pay higher wages even after accounting for workforce composition effects and the other factors in the AKM. Several explanations for this LFWP were discussed by prior literature. We focus on efficiency wages, rent sharing, unionization, and local labor market effects.

The efficiency wage explanation argues that larger firms face higher monitoring complexity and pay higher wages to incentive workers. In Panel A of Table 5, we control for the same proxies for monitoring complexity as in Section 4.4.2. However, neither the fraction of managers per employee, nor the median distance between establishments nor its standard deviation lead to a substantial reduction of the coefficient for total assets. This result indicates that monitoring complexity is not the main driver behind the LFWP.

—Table 6 about here —

Second, larger firms may generate higher rents by exploiting their market power, and they may share some of their higher rents with employees. To measure firms' rents, we use three accounting-based measures: profits per employee, sales per employee, and profits to total assets. When we explore the relation between firm size and these three measures in Figure B.4, we find that profits and sales per employee increase with firm size. For profits to total assets, however, we find lower values for larger firms. When we control for the measures in Panel B of Table 5, none of these proxies has a substantial impact on the coefficient for total assets. Thus, we cannot find any evidence that rent-sharing is the driving force behind the LFWP.

Third, larger firms may have higher unionization rates, which enables employees to extract a higher fraction of rents, and a higher threat of unionization (Dickens and Lang, 1985). Unfortunately, we cannot observe unionization at the firm-level. Therefore, we must measure unionization as the fraction of unionized employees and the fraction of unionized establishments in an industry. To allow for a different effect of industry unionization on larger firms, we include not only industry unionization in our model but also the interaction term with firm size. However, the coefficient estimate on firm size in Panel C of Table 5 ranges in all models from 0.049 to 0.051 which is very close to the baseline estimate of 0.051. Thus, unionization is also unlikely the main reason for the existence of a LFWP.

Fourth, larger firms may cluster in regions with high economic activity and thus a high demand for labor. This can lead to excess demand for labor and increase local wages. To test whether such effects have an impact on the LFWP, we adjust our baseline model and only compare large and small firms in the same region and industry. If the clustering of large firms in particular regions plays an important role for the LFWP, we would expect that the size effect is much weaker after including industry and region fixed effects in the model. The results in Panel D of Table 5 show that the inclusion of industry fixed effects, especially in combination with county fixed effects, reduces the impact of firm size on the LFWP by about one-third (from 0.051 to 0.033).

A potential concern with these results is that firm-level regressions could have limited validity in this context because the county of the firm's headquarter is not necessarily the county in which the majority of the operations takes place. Furthermore, different establishments of the firm may source their workforce from different labor markets, depending on their role in the firm (e.g., R&D center vs. sales office). Thus, we repeat the analysis on the establishment-level in Panel E. The results are even stronger for this specification: the coefficient for firm size drops by about 50% when we include county fixed effects, industry fixed effects, and their interaction. Overall, these results indicate that local effects can help to partially explain the existence of the LFWP.

4.5. Establishment Size within Firms and Wage Inequality

To better understand the role of firms size for wage inequality, we now investigate how size variations of establishments within firms affect wages. Theoretically, two mechanisms can lead to the previously documented higher and more dispersed wages in larger firms. First, the size of the overall firm may matter for wage policies of firms. In this perspective, large firms would pay higher and/or more dispersed wages, independently of the size of a particular establishment within that firm. Second, establishment size may matter for wages, and every establishment of a firm may have its own wage policy, depending on its size. In this perspective, only large establishments within firms would pay higher and/or more dispersed wages. If this holds true, our previous results for larger firms would be driven by the their, on average, larger establishments. Disentangling firm size versus establishment size as potential drivers behind larger firms' higher and more dispersed wages will shed more lights on the reasons for differences in wage policies.

Empirically, we conduct regressions on the establishment-level for this test and include firm fixed effects. This firm fixed effect ensures that the results are driven by size variations of establishments within firms. Because information of total assets is not available on the establishment-level, we use the number of employees in a particular establishment as size proxy in this test. Furthermore, we have to exclude all firms which only operate a single establishment due to the lack of within-firm establishment size variation. Appendix B presents summary statistics at the establishment-level.

4.5.1. Establishment Size within Firms and the Variation of Wages

We first analyze how size differences of establishments within firms affect the variation of wages. For this purpose, we regress the within-establishment variance of the wage and the AKM components on establishment size. Table 7 presents the results. In Panel A, the dependent variable is the within-establishment variance of wages. For comparison, we start by estimating a model without firm fixed effects in Column 1. When we add the firm fixed effect in Column 2, the coefficient estimate for establishment size increases by more than one-third (from 0.0089 to 0.013). To control for industry or region-specific pay policies within a firm, Columns 3 and 4 add interactions of firm fixed effects with establishment-industry and county fixed effects. The coefficients are slightly lower in these specifications (0.012 and 0.010). These results indicate that variations of establishment sizes within firms have an even higher impact on wage inequality than cross-sectional firm-size variations.

— Table 7 about here —

In Panel B, the dependent variable is the within-establishment variance of the person fixed effect. This is the most important AKM component for the within variance of wages on the firm level (cf., Section 4.1). We find that adding a firm fixed effect to the model increases the effect of establishment size on the variation of the person fixed effect (0.011 instead of 0.0072). Adding controls for industry and region reduces the coefficient again, but it still stays above the baseline model. This result is broadly consistent with our prior explanation for the higher variation of workforce quality in larger firms because larger establishments within firms likely have a higher variety of occupations.

The variation of idiosyncratic wage premiums is investigated in Panel C. We find that the firm fixed effect reduces the effect of establishment size (0.0010 instead of 0.0018). Interestingly, the coefficient for establishment size decreases even further when we control for industry (to 0.00088) and region (0.00063). Again, these findings are consistent with the monitoring complexity explanation for higher variance of idiosyncratic wage premiums in larger firms. Although larger establishments are also more difficult to monitor than smaller ones, the main difficulty arises from firm size, not establishment size.

4.5.2. Establishment Size within Firms and the Level of Wages

The next step is to investigate how size differences of establishments within firms affect average wages. For this purpose, we regress the average wage of an establishment and the AKM components on establishment size. Table 8 presents the results. In Panel A, the dependent variable is the average wage. For comparison, we start by estimating a model without firm fixed effects in Column 1. When we add the firm fixed effect in Column 2, the coefficient estimate for establishment size drops substantially (from 0.053 to 0.026). To control for industry or region-specific pay policies within a firm, Columns 3 and 4 add interactions of firm fixed effects with establishment-industry and county fixed effects. While the industry-fixed effect has little impact, the coefficient for establishment size drops to 0.0088 (less than 20% of the baseline coefficient) and becomes statistically insignificant when we only compare small and large establishments in the same region. This result indicates that firm size, not establishment size, is the main driver behind higher wages in larger firms and that regional factors matter a lot for the average establishment wage.

— Table 8 about here —

In panels B, we repeat these analyses for the establishment mean of person fixed effect (i.e., workforce quality). We again find that the inclusion of the firm fixed effect in Column 2 reduces the coefficient for establishment size (from 0.034 to 0.023), but the decrease is less pronounced than for the overall wage in Panel A. When we only compare establishments in the same region in Column 4, the coefficient goes down to 0.0035 and becomes statistically insignificant. This result suggest that firms use workers with a similar quality in the same region, independent of the size of an establishment. Panel C shows the results for the establishment wage premium. The firm fixed effect in Column 2 leads to a substantial reduction of the establishment size coefficient (from 0.022 to 0.0062). Similar to before, the coefficient is even smaller and statistically insignificant after the inclusion of county-fixed effects. This finding shows that the LFWP does not depend on the size of a particular establishment within the firm, at least if regional factors are taken into account. This finding fits well to previously discussed (partial) explanation of the LFWP that larger firms are more likely to be located in regions with high economic activity and a high demand for labour. If a large firm has its headquarter in such a location, it is plausible to assume that the majority of their establishments is located in this region as well, independently of their size. Thus, firm size, not establishment size matters most for the LFWP.

4.6. The Importance of the Firm Structure for Wage Inequality

Our results emphasize the importance of information on firm structures. Without this information, researchers may falsely classify wage heterogeneity across establishments within firms as between inequality. To analyze this effect, we conduct a decomposition of the overall wage inequality into a between and a within component using firm-level and establishment-level information. The procedure in greater details and formulas are presented in Appendix E. Figure 4 presents the results. We find that the average downward bias of the within component to the overall wage variation is about 2.5 percentage points during our sample period from 2010 to 2017.

—Figure 4 about here —

5. Conclusion

In this paper, we analyze the role of firm size for within-firm and between-firm wage inequality. Using a linked employee-establishment-firm data set from Germany which covers about 50,000 individual firms, 120,000 establishments, and twelve million workers, we first find that firm size shows a strong positive correlation to both types of wage inequality.

Second, we decompose wage inequality into workforce composition and non-composition effects. Composition effects are responsible for three-quarters of the relation between firm size and within-firm wage inequality. The corresponding figure for between-firm wage inequality is 50%. Thus, the finding that larger firms pay more unequal and on average higher wages is strongly linked to differences in workforce composition between large and small firms.

Third, we go one step further and try to explain why firm size has a positive impact on the different components of wage inequality. For within-firm wage inequality, we find that higher wage variance in larger firms is largely explainable by more heterogeneous job characteristics and higher employee monitoring complexity. Higher average wages in larger firms are more challenging to explain. Different job characteristics play some role for the higher average workforce quality in larger firms, but they can also not fully explain the large-firm effect. Differences in profitability, monitoring complexity, or unionization levels cannot explain higher wages in larger firms after composition effects (the so-called large-firm wage premium). If anything, local labor markets play some role, but again they can also not fully explain the large-firm effect.

Fourth, we exploit variations of establishment size within firms to shed more light on the relation between (firm) size and wage inequality. Larger establishments of a firm pay more unequal wages, which is mainly driven by their higher variation of workforce quality. However, larger establishments do not pay an economically significant wage premium. This indicates that all establishments of large firms pay a wage premium, relatively independently of their particular size.

Overall, these results show that firms in general and their size in particular matters for wage inequality. However, our findings also suggest that there is little direct (or causal) link from firm size to wage inequality. Rather, differences in wage dispersion and average wages between small and large firms are (at least partly) explainable by factors that are correlated with firm size ("omitted variables"). These are job characteristics, monitoring complexity, and potentially clustering of larger firms in specific regions. However, even in the absence of a direct link from firm size to wage inequality, these results imply that wage inequality will rise if firms in an economy become larger. This is not restricted to the top end of the size distribution (so-called mega firms), but also holds true for size increases of smaller and more representative firms.

References

Abowd, J. M., Kramarz, F., Margolis, D. N., 1999. High wage workers and high wage firms. Econometrica 67 (2), 251–333.

Abowd, J. M., Lemieux, T., 1993. The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada. Quarterly Journal of Economics 108 (4), 983–1014.

Alvarez, J., Benguria, F., Engbom, N., Moser, C., jan 2018. Firms and the Decline in Earnings Inequality in Brazil. American Economic Journal: Macroeconomics 10 (1), 149–189.

Antoni, M., Ganzer, A., vom Berge, P., 2016. Sample of integrated labour market biographies (siab) 1975-2014. FDZ Datenreport. Documentation on Labour Market Data 201604_en, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Antoni, M., Koller, K., Laible, M.-C., Zimmermann, F., 2018. Orbis-adiab: From record linkage key to research dataset: Combining commercial company data with administrative employer-employee data. FDZ Methodenreport 201804_en, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Autor, D., Dorn, D., Katz, L., Patterson, C., Van Reenen, J., may 2019. The Fall of the Labor Share and the Rise of Superstar Firms. NBER Working Paper.

Bloom, N., Guvenen, F., Smith, B. S., Song, J., von Wachter, T. M., 2018. Is the Large Firm Wage Premium Dead or Merely resting? AEA Papers and Proceedings 108, 317–322.

Blossfeld, H.-P., jul 1987. Labor-Market Entry and the Sexual Segregation of Careers in the Federal Republic of Germany. American Journal of Sociology 93 (1), 89–118.

Brown, C., Medoff, J., 1989. The Employer Size-Wage Effect. Journal of Political Economy 97 (5), 1027–1059.

Bryson, A., Erhel, C., Salibekyan, Z., apr 2017. The Effects of Firm Size on Job Quality: A Comparative Study for Britain and France. DoQSS Working Papers 17-08, Department of Quantitative Social Science - UCL Institute of Education, University College London.

Card, D., Cardoso, A. R., Heining, J., Kline, P., 2018. Firms and Labor Market Inequality: Evidence and Some Theory. Journal of Labor Economics 36 (S1), S13–S70.

Card, D., Heining, J., Kline, P., aug 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality*. Quarterly Journal of Economics 128 (3), 967–1015.

Christofides, L. N., Oswald, A. J., 1992. Real Wage Determination and Rent-Sharing in Collective Bargaining Agreements. Quarterly Journal of Economics 107 (3), 985–1002.

Cronqvist, H., Heyman, F., Nilsson, M., Svaleryd, H., Vlachos, J., 2009. Do Entrenched Managers Pay Their Workers More? Journal of Finance 64 (1), 309–339.

Dang, C., (Frank) Li, Z., Yang, C., 2018. Measuring firm size in empirical corporate finance. Journal of Banking and Finance 86, 159–176.

Davis, S. J., Haltiwanger, J., Katz, L. F., Topel, R., 1991. Wage Dispersion between and within U.S. Manufacturing Plants, 1963-86. Brookings Papers on Economic Activity. Microeconomics 1991, 115.

Dickens, W. T., Lang, K., 1985. A test of dual labor market theory. American Economic Review 75 (4), 792–805.

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

Eaton, C., White, W. D., may 1983. The Economy of High Wages: An Agency Problem. Economica 50 (198), 175.

Ellul, A., Pagano, M., Schivardi, F., 2018. Employment and Wage Insurance within Firms: Worldwide Evidence. Review of Financial Studies 31 (4), 1298–1340.

Faggio, G., Salvanes, K. G., van Reenen, J., 2010. The evolution of inequality in productivity and wages: Panel data evidence. Industrial and Corporate Change 19 (6), 1919–1951.

Fitzenberger, B., Osikominu, A., Völter, R., 2006. Imputation rules to improve the education variable in the iab employment subsample. Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften 126 (3), 405–436.

Garen, J. E., aug 1985. Worker Heterogeneity, Job Screening, and Firm Size. Journal of Political Economy 93 (4), 715–739.

Giroud, X., may 2013. Proximity and Investment: Evidence from Plant-Level Data. The Quarterly Journal of Economics 128 (2), 861–915.

Goldschmidt, D., Schmieder, J. F., aug 2017. The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure*. The Quarterly Journal of Economics 132 (3), 1165–1217.

Grullon, G., Larkin, Y., Michaely, R., 2019. Are US Industries Becoming More Concentrated?*. Review of Finance 23 (4), 697–743.

Hamermesh, D. S., jan 1980. Substitution and Labor Market Policy. Challenge 22 (6), 44–47.

Katz, L. F., Autor, D. H., 1999. Chapter 26 Changes in the wage structure and earnings inequality. Handbook of Labor Economics 3 PART (1), 1463–1555.

Kim, H. E., Ouimet, P., 2014. Broad-Based Employee Stock Ownership: Motives and Outcomes. Journal of Finance 69 (3), 1273–1319.

Krueger, A. B., Summers, L. H., 1988. Efficiency Wages and the Inter-Industry Wage Structure. Econometrica 56 (2), 259–293.

Lallemand, T., Plasman, R., Rycx, F., 2005. Why do large firms pay higher wages? Evidence from matched worker-firm data. International Journal of Manpower 26 (7-8), 705–723.

Mueller, H. M., Ouimet, P. P., Simintzi, E., 2017a. Wage inequality and firm growth. American Economic Review 107 (5), 379–383.

Mueller, H. M., Ouimet, P. P., Simintzi, E., 2017b. Within-Firm Pay Inequality. Review of Financial Studies 30 (10), 3605–3635.

Oi, W. Y., 1983. Heterogeneous Firms and the Organization of Production. Economic Inquiry 21 (2), 147–171.

Oi, W. Y., Idson, T. L., 1999. Firm size and wages. In: Handbook of Labor Economics, Amsterdam: North-Holland. pp. 2165–2214.

Pagano, M., Volpin, P., 2005. Managers, Workers, and Corporate Control. Journal of Finance 60 (2), 841–868.

Schmidt, C. M., Zimmermann, K. F., 1991. Work Characteristics, Firm Size and Wages. The Review of Economics and Statistics 73 (4), 705–710.

Shapiro, C., Stiglitz, J. E., 1984. Equilibrium Unemployment as a Worker Discipline Device. The American Economic Review 74 (3), 433–444.

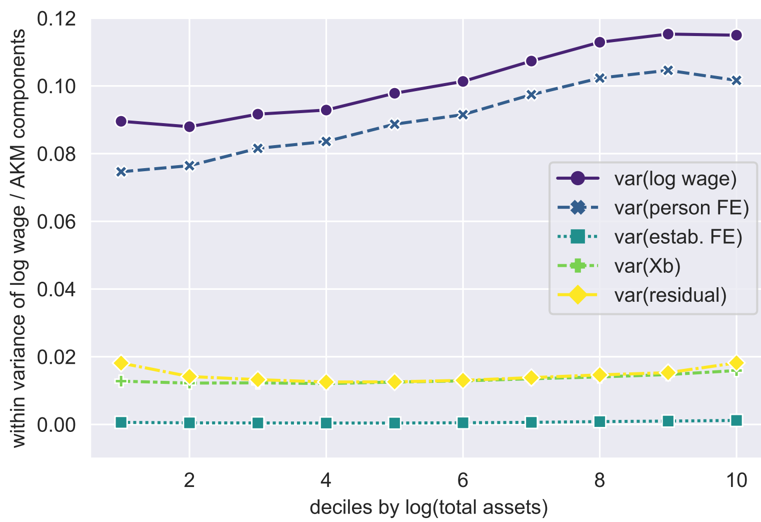
Song, J., Price, D. J., Guvenen, F., Bloom, N., von Wachter, T., feb 2019. Firming Up Inequality*. The Quarterly Journal of Economics 134 (1), 1–50.

Tate, G., Yang, L., 2015. The Bright Side of Corporate Diversification: Evidence from Internal Labor Markets. Review of Financial Studies 28 (8), 2203–2249.

Figures

Figure 1
Wage inequality and firm size
This figure illustrates the relation of wage inequality and firm size. We sort firms into deciles according to total assets. The firm size deciles are plotted on the x-axis. Subfigure (a) presents firm mean wage and its decomposition into AKM components for each decile. Subfigure (b) presents within-firm variance of log wage, person fixed effect, establishment fixed effect, Xb, and residual for each decile. A detailed description of all variables can be found in Appendix A.1.

(a) Wage inequality within firms



(b) Wage inequality between firms

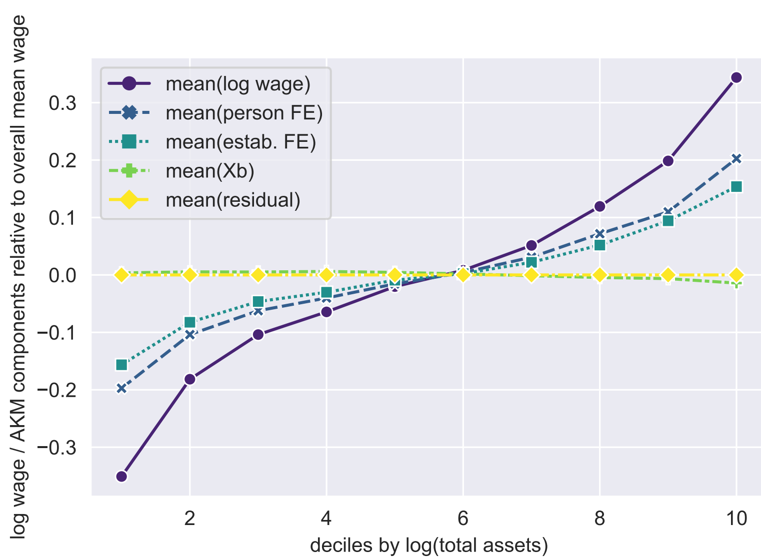
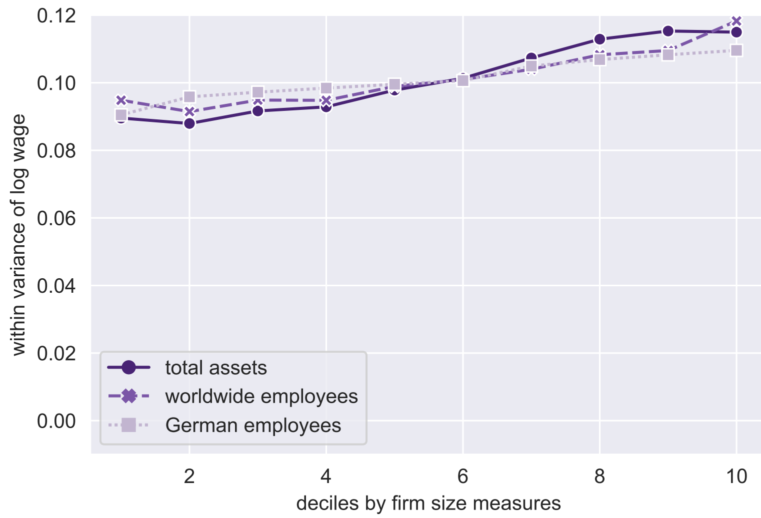


Figure 2

Wage inequality and firm size: capital- and employee-based firm size measures

This figure compares the relation of wage inequality and firm size using capital- and employee-based firm size measures. We sort firms into deciles according to firms' total assets, total German employees and total worldwide employees. The firm size deciles are plotted on the x-axis. Subfigure (a) presents firm mean wage of the deciles. Subfigure (b) presents within-firm variance log wage of deciles. A detailed description of all variables can be found in Appendix A.1.

(a) Wage inequality within firms



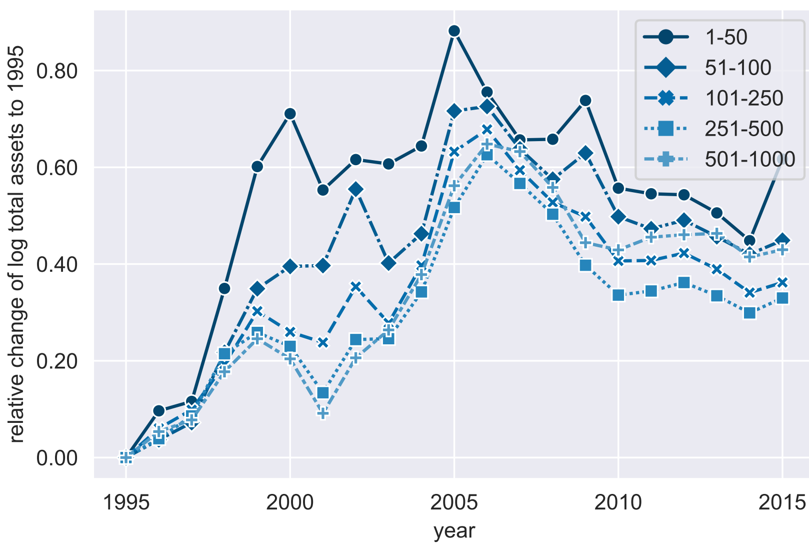
(b) Wage inequality between firms

Figure 3

Development of firm size and wage inequality over the 1995-2015 period

This figure illustrates the development of firm size and wage inequality from 1995 to 2015. We choose this period because accounting information on German firms is very limited before 1995 and we do not (yet) have accounting information on the universe of German firms for the years 2016 and 2017. Subfigure (a) presents the relative change of firm size, which is measured by total assets, compared to 1995. In each year, we distinguish between firms ranked 1 to 50, 51 to 100, 101 to 250, 251 to 500, and 501 to 1000 in terms of firm size. Here, we rely on the universe of German firms from Amadeus. Subfigure (b) presents the within-firm, the between-firm and the overall variation of log wages over time. For this subfigure, we assume that the linking table of establishments to firms, provided by the ORBIS-ADIAB data set (see Section 2.2), is valid in the 1995-2015 period.

(a) Relative change of total assets compared to 1995



(b) Within-firm, between-firm and overall wage inequality

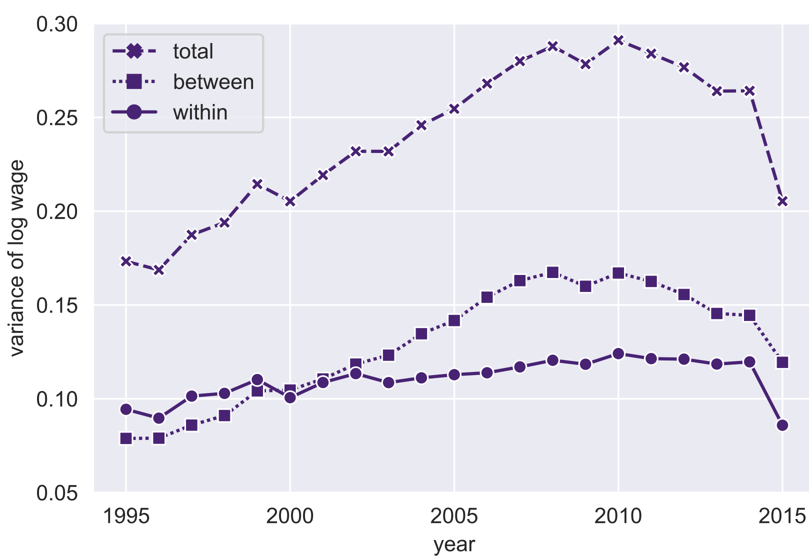
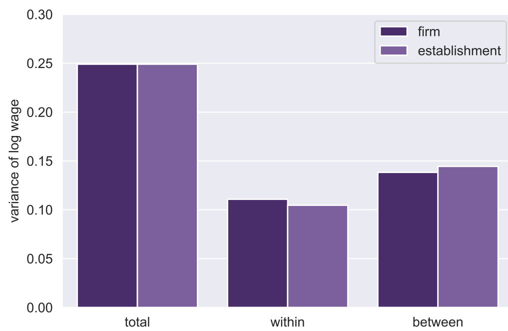


Figure 4

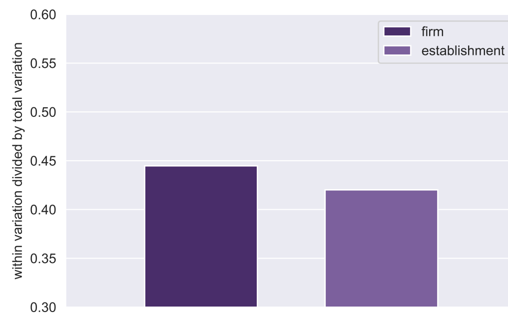
Decomposition of overall wage inequality: firm versus establishment structure

This figure compares the decomposition of overall wage inequality into a between and within component using information on firm and establishment structure in the 2010-2017 period. Subfigure (a) presents the decomposition. Subfigure (b) presents the share of the within component to the overall variation based on firm-level information and on establishment-level information. A detailed description of all variables can be found in Appendix A.1.

(a) Decomposition into between/within



(b) Within / total variation



Tables

Table 1
Descriptive statistics on the firm-level

This table presents descriptive statistics on the firm-level. Reported are the number of observations (Obs), mean value (Mean), standard deviation (SD), median (p50), mean weighted by firms' total German employees (Wgted Mean), and standard deviation weighted by firms' total German employees (Wgted SD). A detailed description of all variables can be found in Appendix A.1.

	Obs	Mean	SD	P50	Wgted Mean	Wgted SD
Panel A: Firm characteristics						
log(total assets)	49,069	15.549	1.622	15.276	17.425	2.005
log(employees$_{firm}$)	49,069	4.220	0.890	3.961	5.774	1.278
log(employees$_{world}$)	45,022	4.446	1.040	4.207	6.027	1.458
sales/employees	45,022	0.233	0.313	0.134	0.192	0.269
profit/employees	28,937	0.022	0.041	0.011	0.017	0.031
profit/total assets	31,053	0.130	0.136	0.115	0.112	0.121
ind. unionization$_{empl}$	49,053	0.388	0.138	0.400	0.410	0.138
ind. unionization$_{estab}$	49,053	0.149	0.089	0.120	0.148	0.090
no. establishments	49,069	1.838	2.827	1.000	6.043	7.741
median(distance)	49,013	0.737	1.596	0.000	2.089	2.332
sd(distance)	49,013	0.474	1.103	0.000	1.040	1.249
manager/employees	49,069	0.050	0.072	0.026	0.061	0.073
complex/employees	49,069	0.177	0.176	0.127	0.197	0.171
highly complex/employees	49,069	0.132	0.175	0.071	0.164	0.170
no. occupations	49,069	21.177	14.187	17.000	43.668	24.296
hhi(occupations)	49,069	0.295	0.203	0.230	0.240	0.195
Panel B: Within-firm variance of log wage and AKM components						
var(log wage)	49,069	0.101	0.051	0.092	0.110	0.049
var(person FE)	49,069	0.090	0.044	0.084	0.095	0.039
var(estab. FE)	49,069	0.001	0.002	0.000	0.002	0.003
var(Xb)	49,069	0.013	0.009	0.011	0.015	0.008
var(residual)	49,069	0.015	0.010	0.012	0.018	0.011
Panel C: Firm mean of log wage and AKM components						
log wage	49,069	4.463	0.328	4.478	4.616	0.367
person FE	49,069	4.188	0.212	4.169	4.274	0.235
firm FE	49,069	0.310	0.164	0.325	0.382	0.176
Xb	49,069	-0.034	0.035	-0.026	-0.041	0.032
residual	49,069	-0.000	0.000	-0.000	-0.000	0.000

Table 2
Wage inequality and firm size

This table presents regressions of wage inequality within and between firms on firm size which is measured by total assets. In panel A, the dependent variables are the within-firm variance of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. In Panel B, the dependent variables are the firm mean of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. For the firm mean of residual, the coefficient on total assets is omitted or virtually zero. This is why we report it as not relevant (n.r.). Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
Panel A: Within-firm inequality					
	var(log wage)	var(person FE)	var(estab. FE)	var(Xb)	var(residual)
log(total assets)	0.0045***	0.0034***	0.00033***	0.00056***	0.00088***
	(7.63)	(8.28)	(6.02)	(7.84)	(5.84)
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.03	0.03	0.04	0.02	0.03
Panel B: Between-firm inequality					
	log wage	person FE	firm FE	Xb	residual
log(total assets)	0.12***	0.067***	0.051***	-0.0022***	n.r.
	(20.13)	(17.12)	(19.34)	(-8.11)	n.r.
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.40	0.33	0.34	0.02	n.r.

Table 3
Within-firm inequality: explaining worker heterogeneity in larger firms

The dependent variable is the within-firm variance of person fixed effect. Column 1 presents the baseline model regressing the variance of person fixed effect on total assets. In columns 2 to 5, we add control variables for the dispersion of occupations within firms to the baseline model. These are the number of employees performing complex tasks standardized by total German employees, the number of employees performing highly complex tasks standardized by total German employees, and the Herfindahl index as concentration measure of occupations. Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
log(total assets)	0.0034***	0.0026***	0.0018***	0.0023***	0.00075*
	(8.28)	(6.89)	(5.06)	(5.75)	(1.75)
complex/employees		0.040***			0.021***
		(10.39)			(6.01)
highly complex/employees			0.072***		0.064***
			(20.23)		(17.27)
hhi(occupations)				-0.032***	-0.024***
				(-10.39)	(-6.22)
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.03	0.06	0.12	0.05	0.15

Table 4

Within-firm inequality: explaining greater variation of idiosyncratic wage premium in larger firms - measured excluding managers

The dependent variable is the within-firm variance of residual excluding employees with the occupational code 751 Entrepreneur, Manager and Division Manager. Column 1 presents the baseline model regressing the variance of residuals on total assets. In columns 2 to 5, we add control variables for monitoring complexity to the baseline model. These are the number of managers employed divided by German employees, the median distance between a firm's headquarter and establishments, and the standard deviation of this distance. Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
log(total assets)	0.00077***	0.00038**	0.00025**	0.00053***	-0.000061
	(5.30)	(2.45)	(2.41)	(3.86)	(-0.55)
manager/employees		0.040***			0.037***
		(10.40)			(10.50)
median(distance)			0.0010***		0.00087***
			(9.34)		(4.78)
sd(distance)				0.0012***	0.00014
				(8.40)	(0.50)
Obs	49,069	49,069	49,013	49,013	49,013
R2	0.02	0.09	0.06	0.04	0.13

Table 5

Between-firm inequality: explaining higher worker quality in larger firms

The dependent variable is the firm mean of person fixed effect. Column 1 presents the baseline model regressing the mean person fixed effect on total assets. In columns 2 to 5, we add control variables for the complexity of occupations to the baseline model. These are the number of employees performing complex tasks standardized by total German employees, the number of employees performing highly complex tasks standardized by total German employees, and the Herfindahl index as concentration measure of occupations. Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
log(total assets)	0.067***	0.056***	0.052***	0.062***	0.044***
	(17.12)	(14.00)	(12.46)	(14.55)	(11.75)
complex/employees		0.58***			0.42***
		(17.35)			(17.97)
highly complex/employees			0.70***		0.58***
			(20.77)		(21.50)
hhi(occupations)				-0.15***	-0.058**
				(-3.43)	(-2.24)
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.33	0.50	0.56	0.34	0.65

Table 6
Between-firm inequality: explaining large firm wage premium

The dependent variable is the firm mean of establishment fixed effect. Column 1 presents the baseline model regressing the mean establishment fixed effect on total assets. In columns 2 to 5, we add control variables and fixed effects to the baseline model. In panel A, we add control variables for monitoring complexity. These are the managers per employees, the median distance between firm headquarter and establishments, and the standard deviation of the distance. In panel B, we include measures for firms' rent. These are the profit per employees, the sales per employees, and the profit standardized by total assets. In panel C, we control for industry unionization. In detail, we add the industry unionization rate based on employees, its interaction term with total assets, the industry unionization rate based on establishments, and its interaction term with total assets to the model. For the interaction terms, the variables are centered at the mean. In Panel D, we add industry, county, and industry-county fixed effects to control for local effects. Panels A to D are based on the firm-level data set and regressions are weighted by firms' total number of employees. In Panel E, we test the role of local effects on establihment level. This allows us to control for the industry of the establishment, the county of the establishment, and the combination of establishment industry and county. Establishment-level regressions are weighted by establishments' number of employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
Panel A: Efficiency wages					
log(total assets)	0.051***	0.048***	0.058***	0.051***	0.054***
	(19.34)	(16.32)	(32.80)	(18.18)	(27.95)
manager/employees		0.36***			0.39***
		(5.39)			(5.90)
median(distance)			-0.012***		-0.020***
			(-5.09)		(-5.68)
sd(distance)				0.00037	0.020***
				(0.21)	(4.52)
Obs	49,069	49,069	49,013	49,013	49,013
R2	0.34	0.37	0.37	0.35	0.40
Panel B: Rent sharing					
log(total assets)	0.051***	0.049***	0.051***	0.051***	0.050***
	(19.34)	(12.94)	(18.51)	(14.62)	(12.89)
profit/employees		0.49***			0.13
		(4.09)			(1.11)
sales/employees			0.070***		0.068***
			(5.34)		(7.21)
profit/total assets				0.023	0.00048
				(0.91)	(0.01)
Obs	49,069	28,937	45,022	31,053	28,937
R2	0.34	0.27	0.35	0.27	0.27

Continued on next page

Table 6 continued

Panel C: Industry unionization					
log(total assets)	0.051***	0.051***	0.049***	0.051***	0.050***
	(19.34)	(19.68)	(21.87)	(18.58)	(19.09)
ind. unionization$_{empl}$		-0.053	-0.17***		
		(-1.51)	(-9.64)		
log(total assets) x unionization$_{empl}$			0.065***		
			(3.57)		
ind. unionization$_{estab}$				-0.18***	-0.29***
				(-3.76)	(-11.05)
log(total assets) x unionization$_{estab}$					0.052**
					(2.37)
Obs	49,069	49,053	49,053	49,053	49,053
R2	0.34	0.35	0.36	0.35	0.36
Panel D: Local effects on firm-level					
log(total assets)	0.051***	0.039***	0.049***	0.033***	0.033***
	(19.34)	(18.84)	(39.08)	(34.35)	(35.46)
Industry FE	No	Yes	No	Yes	Yes
County FE	No	No	Yes	Yes	Yes
Industry x county FE	No	No	No	No	Yes
Obs	49,069	49,035	49,030	48,998	35,935
R2	0.34	0.58	0.49	0.69	0.83
Panel E: Local effects on establishment-level					
log(total assets)	0.052***	0.035***	0.048***	0.032***	0.026***
	(19.42)	(12.99)	(18.62)	(14.06)	(7.44)
Estab. industry FE	No	Yes	No	Yes	Yes
Estab. county FE	No	No	Yes	Yes	Yes
Estab. ind. x county FE	No	No	No	No	Yes
Obs	119,641	119,598	117,912	117,880	88,558
R2	0.32	0.59	0.45	0.69	0.83

Table 7

Wage inequality within establishments and establishment size

The dependent variables are the within-establishment vairance of log wage in Panel A, within-establishment variance of person fixed effect in Panel B, and within-establishment variance of residual in Panel C. In column 1, the dependent variables are regressed on establishments' number of employees. In column 2, we add firm fixed effects to the model to compare only establishments within firms. In columns 3, we include firm fixed effects interacted with establishment industry fixed effects to compare establishments within firms and in the same industry. In columns 4, we include firm fixed effects interacted with establishment county fixed effects to compare establishments within firms and in the same county. The sample is constructed on establishment level and only consists of establishments that belong to a firm with at least two establishments. Regressions are weighted by establishments' number of employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)
Panel A: Variance(log wage)				
log(employees$_{estab}$)	0.0089***	0.013***	0.012***	0.010***
	(9.31)	(24.16)	(19.19)	(6.26)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	79,770	79,587	67,956	29,859
R2	0.04	0.64	0.72	0.78
Panel B: Variance(person FE)				
log(employees$_{estab}$)	0.0072***	0.011***	0.0097***	0.0081***
	(10.96)	(27.81)	(20.96)	(12.20)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	79,764	79,580	67,951	29,857
R2	0.06	0.68	0.72	0.81
Panel C: Variance(residual)				
log(employees$_{estab}$)	0.0018***	0.0010***	0.00088***	0.00063**
	(9.74)	(7.97)	(6.74)	(2.16)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	79,764	79,580	67,951	29,857
R2	0.07	0.69	0.75	0.83

Table 8

Wage inequality between establishments and establishment size

The dependent variables are the establishment mean of log wage in Panel A, person fixed effect in Panel B, and establishment fixed effect in Panel C. In column 1, the dependent variables are regressed on establishments' number of employees. In column 2, we add firm fixed effects to the model to compare only establishments within firms. In columns 3, we include firm fixed effects interacted with establishment industry fixed effects to compare establishments within firms and in the same industry. In columns 4, we include firm fixed effects interacted with establishment county fixed effects to compare establishments within firms and in the same county. The sample is constructed on establishment level and only consists of establishments that belong to a firm with at least two establishments. Regressions are weighted by establishments' number of employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)
Panel A: Log wage				
log(employees$_{estab}$)	0.053***	0.026***	0.025***	0.0088
	(6.15)	(13.87)	(11.48)	(1.12)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	81,840	81,840	69,998	30,932
R2	0.42	0.93	0.95	0.96
Panel B: Person FE				
log(employees$_{estab}$)	0.034***	0.023***	0.021***	0.0035
	(5.26)	(12.02)	(8.89)	(0.68)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	81,840	81,840	69,998	30,932
R2	0.30	0.86	0.90	0.93
Panel C: Establishment FE				
log(employees$_{estab}$)	0.022***	0.0062***	0.0060***	0.0054
	(6.77)	(5.43)	(4.16)	(1.59)
Firm FE	No	Yes	Yes	Yes
Firm FE x Estab. industry FE	No	No	Yes	No
Firm FE x Estab. county FE	No	No	No	Yes
Obs	81,840	81,840	69,998	30,932
R2	0.37	0.91	0.92	0.95

Appendix A. Definition of Variables

Table A.1
Definition of Variables

Variable	**Description**
Wage and AKM componentens	
log wage	Imputed real log daily wage. Source: IEB (Integrated Employment Biographies).
person FE	Person fixed effect from the AKM-type regression. The implementation and interpretation of the AKM-type regression is explained in detail in Section 3.3.
establishment FE	Establishment fixed effect from the AKM-type regression. The implementation and interpretation of the AKM-type regression is explained in detail in Section 3.3.
firm FE	Employee-weighted average of a firm's establishment fixed effects from the AKM-type regression.
Xb	Combination of life cycle and aggregate factors from the AKM-type regression. The implementation and interpretation of the AKM-type regression is explained in detail in Section 3.3.
residual	Residual from the AKM-type regression. The implementation and interpretation of the AKM-type regression is explained in detail in Section 3.3.
Firm / establishment size	
log(total assets)	Natural logarithm of total assets (toas) in m of 2013 EUR. Source: Orbis.
log(employees$_{firm}$)	Natural logarithm of the firm's total employees in Germany. Source: BHP (Betriebshistorik Panel), IAB.
log(employees$_{world}$)	Natural logarithm of the firm's worldwide employees. Source: Orbis.
log(employees$_{estab}$)	Natural logarithm of the establishment's employees. Source: BHP, IEB.
Other firm characteristics	
sales/employees	Total sales standardized by total number of worldwide employees in m of 2013 EUR per employee ($\frac{sales}{employees_{world}}$). Source: Orbis.
profit/employees	Earnings before interest, taxes, depreciation, and amortization standardized by total number of worldwide employees in m of 2013 EUR per employee ($\frac{ebta}{employees_{world}}$). Source: Orbis.
profit/total assets	Earnings before interest, taxes, depreciation, and amortization standardized by total assets ($\frac{ebta}{toas}$). Source: Orbis.
no. establishments	Firm's number of establishments. Source: Oribs-ADIAB.

continued on next page

appendix A.1 continued

Variable	Description
median(distance)	Median of logarithm of distance (in kilometers) from county of establishment to county of firm headquarter (ln(distance + 1)). If establishment and headquarter are located in same county, the distance is zero. Source: Own calculation.
sd(distance)	Standard deviation of logarithm of distance (in kilometers) from county of establishment to county of firm headquarter (ln(distance + 1)). If all establishment are in same county as headquarter or the firm has only one establishment, it set to zero. Source: Own calculation.
manager/employees	Number of employed managers (according to (Blossfeld, 1987)) divided by total employees. Source: BHP, IEB.
complex/employees	Number of employees who perform complex tasks dividend by total employees. Source: BHP, IEB.
highly complex/employees	Number of employees who perform highly complex tasks dividend by total employees. Source: BHP, IEB.
no. occupations	Number of occupations according to the first three digits of the Classification of Occupations 1988 (KldB 1988). Source: IEB.
hhi(occupations)	Herfindahl index as concentration measure of employees' occupations according to the first three digits of the Classification of Occupations 1988 (KldB 1988). Source: IEB.
Establishment and firm characteristics used to construct fixed effects	
county$_{firm}$	County ("Landkreis") of firm's headquarter. Source: Orbis.
county$_{estab}$	County ("Landkreis") of establishment. Source: BHP.
industry$_{firm}$	3-digits of the classifcation of economic activities WZ 2008. Source: Orbis.
industry$_{estab}$	3-digits of the classifcation of economic activities WZ 2008. Source: BHP.

BHP stands for Betriebshistorik Panel provided by the Institute of Employment Research, IEB for Integrated Employment Biographies provided by the Institute of Employment Research, and Orbis for the Orbis database by Bureau van Dijk.

Appendix B. Further Descriptive Statistics

Table B.1
Descriptive statistics on the establishment-level

This table presents descriptive statistics on the establishment-level. Reported are the number of observations (Obs), mean value (Mean), standard deviation (SD), median (p50), mean weighted by establishments' number of employees (Wgted Mean), and standard deviation weighted by establishments' number of employees (Wgted SD). A detailed description of all variables can be found in Appendix A.1.

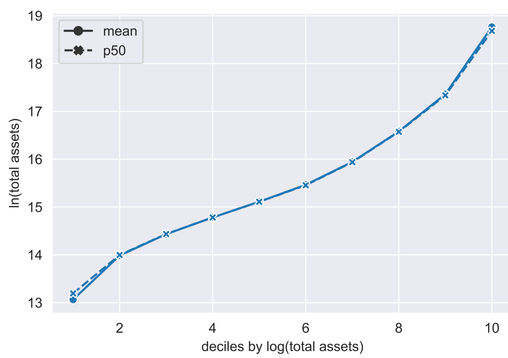
	Obs	Mean	SD	P50	Wgted Mean	Wgted SD
Panel A: Firm and establishment characteristics						
log(total assets)	119,641	16.755	2.021	16.728	17.422	2.004
log(employees$_{estab}$)	119,641	2.993	1.443	3.234	5.098	1.200
log(employees$_{firm}$)	119,641	5.242	1.374	5.118	5.781	1.274
log(employees$_{world}$)	113,559	5.621	1.571	5.559	6.032	1.455
distance	117,854	2.703	2.586	3.106	1.377	2.260
(complex/employees)$_{estab}$	116,010	0.180	0.245	0.091	0.193	0.184
(h. complex/employees)$_{estab}$	116,010	0.162	0.244	0.062	0.160	0.182
Panel B: Within-establishment variance of log wage and AKM components						
var$_{estab}$(log wage)	117,565	0.086	0.062	0.076	0.104	0.054
var$_{estab}$(person FE)	117,565	0.074	0.052	0.066	0.090	0.042
var$_{estab}$(Xb)	117,565	0.012	0.011	0.009	0.015	0.009
var$_{estab}$(residual)	117,565	0.015	0.014	0.011	0.018	0.012
Panel C: Establishment mean of log wage and AKM components						
log wage$_{estab}$	119,641	4.457	0.365	4.461	4.612	0.381
person FE$_{estab}$	119,641	4.183	0.253	4.158	4.272	0.251
estab. FE$_{estab}$	119,641	0.311	0.189	0.326	0.381	0.184
Xb$_{estab}$	119,641	-0.036	0.049	-0.027	-0.041	0.038
residual$_{estab}$	119,641	0.000	0.000	-0.000	-0.000	0.000

Figure B.1

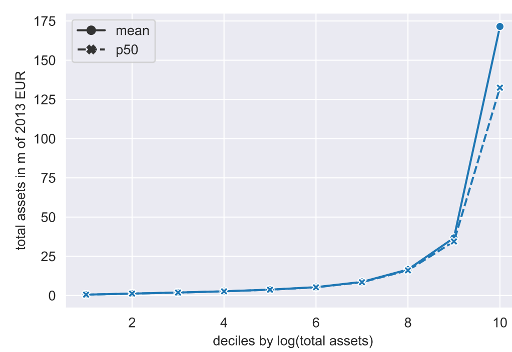
Differences between smaller and larger firms: firm size

We sort firms into deciles according to total assets. The firm size deciles are plotted on the x-axis. The variable on the y-axis is stated in the title of each subfigure. A detailed description of all variables can be found in Appendix A.1.

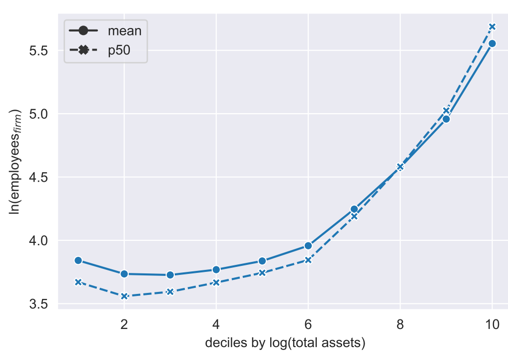
(a) Log of total assets



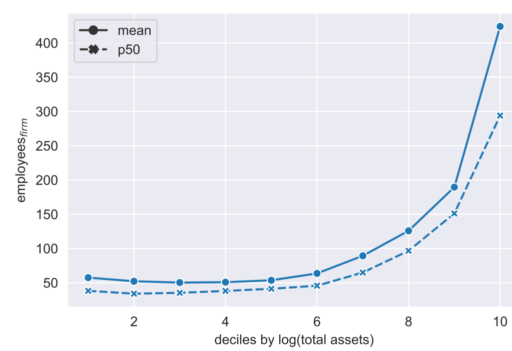
(b) Total assets in m of 2013 EUR



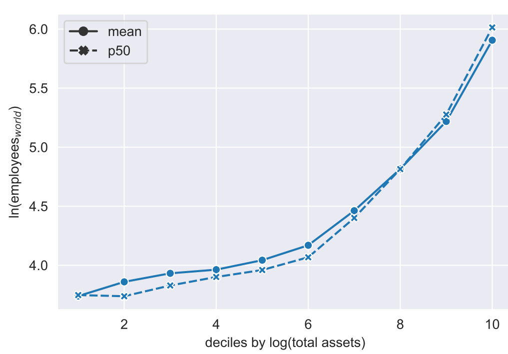
(c) Log of German employees



(d) German employees



(e) Log of worldwide employees



(f) Worldwide employees

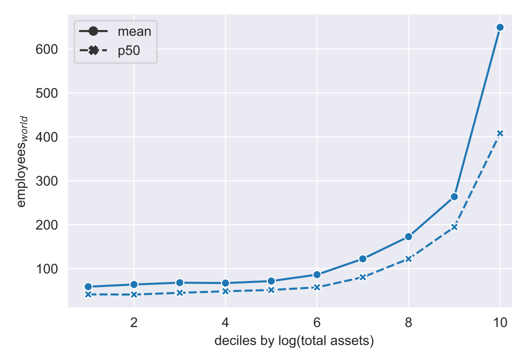
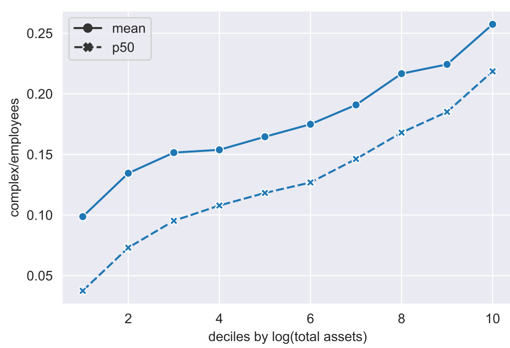


Figure B.2

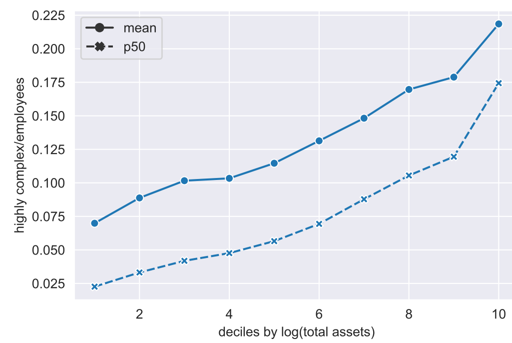
Differences between smaller and larger firms: job characteristics

We sort firms into deciles according to total assets. The firm size deciles are plotted on the x-axis. The variable on the y-axis is stated in the title of each subfigure. A detailed description of all variables can be found in Appendix A.1.

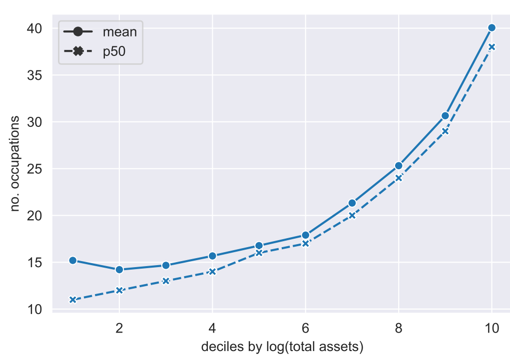
(a) Complex/employees



(b) Highly complex/employees



(c) Number of occupations



(d) Herfindahl index of occupations

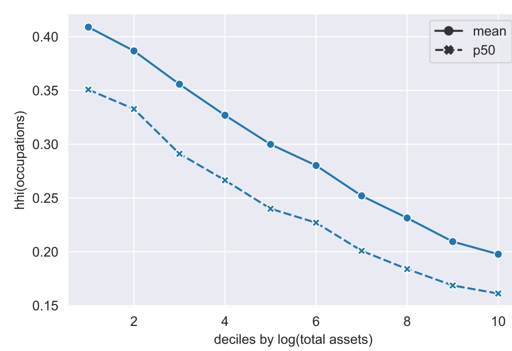
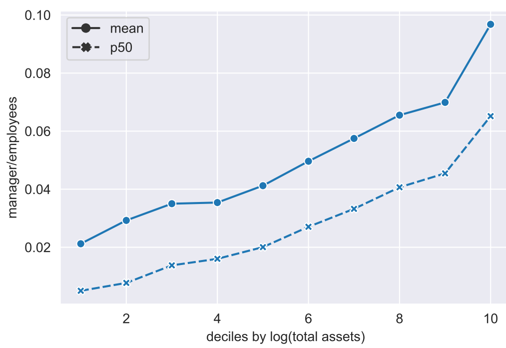
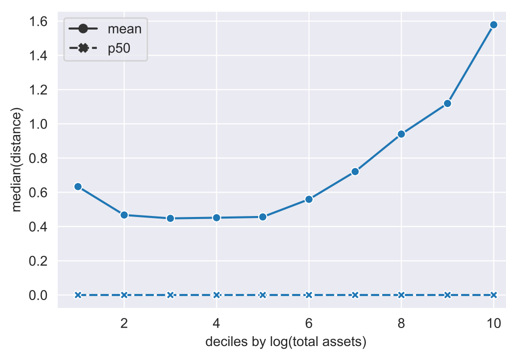


Figure B.3
Differences between smaller and larger firms: monitoring complexity
We sort firms into deciles according to total assets. The firm size deciles are plotted on the x-axis. The variable on the y-axis is stated in the title of each subfigure. A detailed description of all variables can be found in Appendix A.1.

(a) Manger/employees



(b) Median(distance)



(c) Sd(distance)

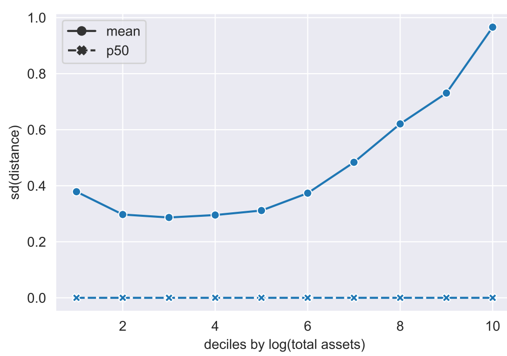
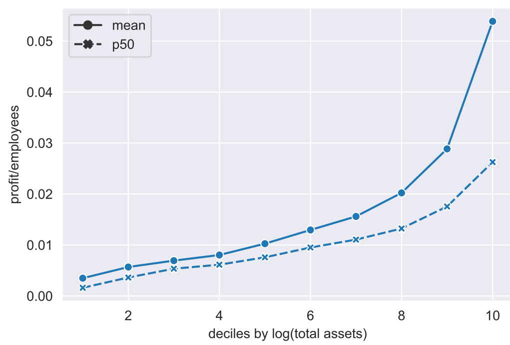
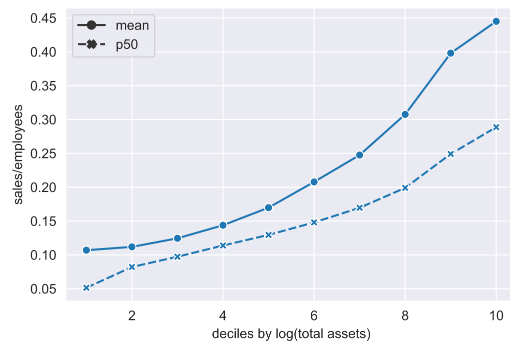


Figure B.4
Differences between smaller and larger firms: rent sharing
We sort firms into deciles according to total assets. The firm size deciles are plotted on the x-axis. The variable on the y-axis is stated in the title of each subfigure. A detailed description of all variables can be found in Appendix A.1.

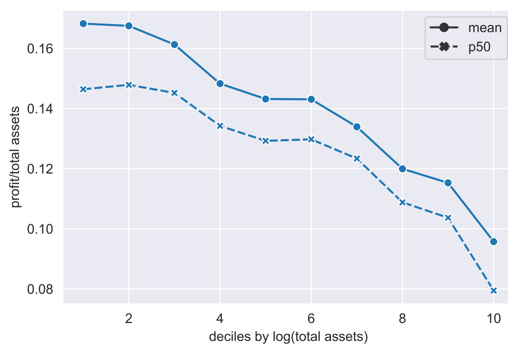
(a) Profit/employees in m of EUR 2013



(b) Sales/employees in m of EUR 2013



(c) Profit/total assets



Appendix C. Alternative Measurement of Firm Size

Table C.1
Wage inequality and firm size: measured by worldwide employees

This table presents regressions of wage inequality within and between firms on firm size which is measured by total number of worldwide employees. In panel A, the dependent variables are the within-firm variance of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. In Panel B, the dependent variables are the firm mean of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. For the firm mean of residual, the coefficient on total assets is omitted or virtually zero. This is why we report it as not relevant (n.r.). Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
Panel A: Within-firm inequality					
	var(log wage)	var(person FE)	var(estab. FE)	var(Xb)	var(residual)
log(empl$_{world}$)	0.0055***	0.0032***	0.00064***	0.00081***	0.0016***
	(6.19)	(5.25)	(8.07)	(7.20)	(7.01)
Obs	45,022	45,022	45,022	45,022	45,022
R2	0.03	0.01	0.09	0.02	0.05
Panel B: Between-firm inequality					
	log wage	person FE	firm FE	Xb	residual
log(empl$_{world}$)	0.096***	0.056***	0.043***	-0.0027***	n.r.
	(10.46)	(8.78)	(11.01)	(-6.22)	n.r.
Obs	45,022	45,022	45,022	45,022	45,022
R2	0.15	0.12	0.13	0.01	n.r.

Table C.2

Wage inequality and firm size: measured by total German employees

This table presents regressions of wage inequality within and between firms on firm size which is measured by total number of German employees. In panel A, the dependent variables are the within-firm variance of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. In Panel B, the dependent variables are the firm mean of log wage, person fixed effect, establishment fixed effect, Xb and residual from the AKM-type regression. For the firm mean of residual, the coefficient on total assets is omitted or virtually zero. This is why we report it as not relevant (n.r.). Regressions are weighted by firms' total number of German employees. T-statistics based on Huber/White robust standard errors clustered by firms are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.1.

	(1)	(2)	(3)	(4)	(5)
Panel A: Within-firm inequality					
	var(log wage)	var(person FE)	var(estab. FE)	var(Xb)	var(residual)
log(empl$_{firm}$)	0.0055***	0.0025***	0.00074***	0.0012***	0.0022***
	(5.73)	(3.74)	(8.67)	(10.66)	(9.99)
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.02	0.01	0.09	0.03	0.07
Panel B: Between-firm inequality					
	log wage	person FE	firm FE	Xb	residual
log(empl$_{firm}$)	0.11***	0.061***	0.049***	-0.0036***	n.r.
	(10.72)	(8.81)	(11.81)	(-8.09)	n.r.
Obs	49,069	49,069	49,069	49,069	49,069
R2	0.14	0.11	0.13	0.02	n.r.

Appendix D. Decomposition of Within-firm and Within-establishment Variances

Table D.1

Full decomposition of within-firm and within-establishments variances

This table presents the full decomposition of the within variance on firm- and establishment-level into the AKM components. Panel A presents the decomposition of the within-firm variance. Panel B presents the decomposition of the wihin-establishment variance. Presented are the employee-weighted mean values of within variances and covariances. A detailed description of all variables can be found in Appendix A.1.

	Obs	Var Component	Share
Panel A: Within-firm variance			
var(log wage)	49,069	0.110	1.000
var(person FE)	49,069	0.095	0.865
var(estab. FE)	49,069	0.002	0.015
var(Xb)	49,069	0.015	0.139
var(residual)	49,069	0.018	0.161
2cov(person FE, Xb)	49,069	-0.020	-0.178
2cov(person FE, residual)	49,069	0.000	0.002
2cov(Xb, residual)	49,069	-0.000	-0.004
2cov(person FE, estab. FE)	49,069	-0.000	-0.002
2cov(estab. FE, Xb)	49,069	0.000	0.000
2cov(estab. FE, residual)	49,069	0.000	0.000
Panel B: Within-establishment variance			
var_{estab}(log wage)	117,565	0.104	1.000
var_{estab}(person FE)	117,565	0.090	0.864
var_{estab}(Xb)	117,565	0.015	0.145
var_{estab}(residual)	117,565	0.018	0.173
2cov_{estab}(person FE, Xb)	117,565	-0.018	-0.177
2cov_{estab}(person FE, residual)	117,565	0.000	0.001
2cov_{estab}(Xb, residual)	117,565	-0.000	-0.003

Appendix E. Decomposition of Overall Wage Inequality into a Between and a Within Component

We decompose the overall variance of log daily wage into between-firm and within-firm variation. As we observe information on firms and establishments, we can also compare this to the decomposition into a between-establishment and a within-establishment component ignoring firm-level information.

The overall variance of wage can be decomposed into a between- and a within-firm component,

$$var(y_t^{i,j,k}) = var(\bar{y}_t^k) + \sum_k w_k \times var_k(y_t^{i,j,k} | i \in j, j \in k), \tag{E.1}$$

where $var(\bar{y}_t^k)$ is the between-firm variance of firm mean wage, and the second term is the employment-weighted mean of within-firm variance in employee wage. w_k denotes the employment share of firm k in the sample. Alternatively, we can decompose the overall variance into a between- and within-establishment component ignoring the firm-level information,

$$var(y_t^{i,j,k}) = var(\bar{y}_t^j) + \sum_j w_j \times var_j(y_t^{i,j,k} | i \in j) \tag{E.2}$$

This allows us to explore the difference in the share of within variation to overall variation using firm and establishment information,

$$\frac{\sum_j w_j \times var_j(y_t^{i,j,k} | i \in j)}{var(y_t^{i,j,k})} - \frac{\sum_k w_k \times var_k(y_t^{i,j,k} | i \in j, j \in k)}{var(y_t^{i,j,k})} \tag{E.3}$$

The comparison is interesting because the identification of firm structures was not possible in German administrative data until the release of ORBIS-ADIAB data set (see Section 2.2) and gains importance given the increasing complexity of firm structures.