Why do large firms pay higher wages? Novel stylized facts from linked firm-establishment-worker data☆

Daniel Bias[a], Chen Lin[b], Ben 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

Little is known why the large firm wage premium (LFWP) exists. Using a matching of administrative employee-level information from Germany with firm-level data allows us to analyze how size affects the wage premium after adjusting for worker characteristics. We find that (i) firm size, not establishment size matters for the LFWP, (ii) the premium is driven by assets, not employees, (iii) the premium is a concave function of firm size, (iv) firm profitability, market share, and productivity have little impact on the premium's magnitude, (v) time-constant omitted variables explain about one-third of the premium, and (vi) reverse causality is unlikely to cause the LFWP. These findings are consistent with the idea that firms pay efficiency wages to incentivize workers but fail to support rent-sharing explanations.

Keywords: Large-firm wage premium, efficiency wages, rent sharing

☆We thank Christoph Kaserer and Christian Merkl for comments and suggestions. Please send correspondence to Thomas Schmid, E-Mail: schmid@hku.hk

November 1, 2018

1. Introduction

Wage differences across firms exist even after controlling for workforce composition and worker characteristics (e.g., Card, Heining and Kline, 2013; Song et al., 2018). Those firm-specific premiums explain about 20% of the overall wage variation. It has been shown that firm size has a strong explanatory power for wage differentials between firms (e.g., Brown and Medoff, 1989; Oi and Idson, 1999; Bloom et al., 2018). Although the existence of this large-firm wage premium (LFWP)[1] is well documented, surprisingly little is known about the reasons why larger firms pay wage premiums.

One challenge that the empirical literature on firm-specific wage premiums typically faces is to link data on workers' wages and characteristics to firm-level data sets. Employee-level data is only available on an establishment-level, not on a firm level, in many settings. However, firm-level data is necessary to shed light on the reasons why the LFWP exists because wages are ultimately decided at the firm-level. For instance, the importance of firms for wage setting was recently documented by Gabaix and Landier (2008), Mueller, Ouimet and Simintzi (2017a), or Mueller, Ouimet and Simintzi (2017b). Furthermore, firm capital as size measure is only available at a firm-level. The fall of the labor share in the largest firms (Autor et al., 2017) underlines the importance to differentiate between firm size based on capital and labor.

In this paper, we use a recently conducted match of administrative employee-level information from Germany with firm-level data from ORBIS. This dataset contains 218,260 establishments that belong to 169,173 firms in the period 2010 to 2016. For some tests, we additionally use data back until 1993. This dataset has three major advantages. First, the detailed employee-level wage data enable us to calculate the wage premium after controlling for worker characteristics in the regression framework of Abowd, Kramarz and Margolis (1999), henceforth AKM.[2] Controlling for worker characteristics is important

[1]Throughout this paper, we define the large firm wage premium (LFWP) as the difference between the average wages of workers in large and small firms, after controlling for observable and unobservable worker characteristics. Most early studies in this topic examine the difference in average wages between large and small firms without controlling for observed and unobserved worker characteristics. Oi and Idson (1999) provide a detailed survey of this large literature.

[2]We apply the AKM implementation by Card, Heining and Kline (2013) for the German labor market. The fixed-effects regression framework of AKM allows us to estimate a separate fixed effect for each worker and each firm. We interpret the firm fixed effect as the

because the workforce composition may depend on firm size and higher wages in larger firms may then simply reflect differences in workers' characteristics. Ouimet and Zarutskie (2014), for instance, document that younger and thus smaller firms employ a disproportionately high fraction of young employees. More generally, AKM show that about 75% of the firm-size wage effect can be explained by worker characteristics. Second, the dataset allows to link single establishments to firms in order to create a matched firm-establishment data set. This is crucial because wages are typically determined on the firm, not the establishment level. Third, the new link of the employee-level data to ORBIS enables us to measure firm-level characteristics such as total assets, profitability, or market share. This novel dataset will allow us to provide new evidence on the characteristics of the LFWP and link them to possible explanations.

Two common theoretical explanations for the LFWP are rent sharing and efficiency wages.[3] Rent sharing can explain higher wages in large firms if large firms generate more rents and share the same fraction of rents with workers than smaller firms. Larger firms may generate higher rents because larger firms can exploit their market power to maximize profits (e.g., Christofides and Oswald, 1992; Abowd and Lemieux, 1993). Alternatively, rent sharing as stand-alone factor can also explain the LFWP if large firms generate similar rents than smaller firms, but share a larger fraction with their workers. The sharing function is typically the outcome of the bargaining process between employees and the firm and thus influenced by the power of employees (e.g., unionized vs non-unionized workers). Thus, if workers have on average more power in larger firms (e.g., because larger firms are more likely unionized), larger firms may agree to share a higher fraction of rents with their workers.[4]

Efficiency wage theories explain that firms pay non-competitive wages to minimize turnover cost, to improve the ex-ante selection of workers, or to in-

premium that a firm pays to a typical worker relative to the average firm (after controlling for observed and unobserved worker characteristics).

[3]Although we focus on these two theories because they belong to the most popular explanations of the LFWP, others exist as well. Among those are large versus small firm differences in 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).

[4]A similar mechanism has been used as explanation for excessive CEO pay. Among others, Bebchuk and Fried (2004) and Bebchuk, Cremers and Peyer (2011) argue that excessive CEO compensation is the outcome of rent extraction by powerful CEOs.

centivize workers ex post (Krueger and Summers, 1988). For turnover cost, the basic idea is firms are negatively affected by employee turnover. Higher pay may reduce turnover probability and thus avoid adverse effects for the firm. If turnover cost or turnover probability are higher for larger firms, they may offer higher wages to workers. For the ex-ante selection effect, offering a higher wage may attract a larger and higher quality pool of applicants. This is beneficial for the firm because workers' quality is not perfectly observable. If large firms would face more problems in attracting (talented) workers or in observing the quality of workers ex-ante than smaller firms, they may choose to offer higher wages. Lastly, higher wages may have an incentive effect for incumbent workers. Because monitoring is costly, the firm cannot perfectly observe workers' efforts. Higher wages increases workers increases their opportunity cost and makes loosing their job more costly for them. Because monitoring is likely more costly for large firms, they may pay higher wages relative to small firms.

Empirically, we first show that—in line with previous literature—the wage premium increases in firm size. Next, we establish several novel stylized facts of the LFWP and we compare those stylized facts to the predictions of rent sharing and efficiency wages theories. The dependent variable in our models is the establishment-level wage premium obtained by the AKM model. Because the AKM methodology requires several years to estimate the wage premium, we use the period from 2010 to 2016 to estimate the wage premium in our main analyses. Our main models include only one observation per establishment (i.e, the average wage premium between 2010 and 2016) and thus exploit cross-sectional heterogeneity in firm size.

The first stylized fact that we identify is that firm size, not establishment size matters for the LFWP. The effect of establishment size on the wage premium becomes economically insignificant after we control for firm-fixed effects. This indicates that the LFWP is a firm-level premium which is paid in all establishments of a firm, independently of their size.[5]

The second stylized fact is that the LFWP is driven by firms' total assets, not their number of employees. In particular, the effect of total employees becomes negative but economically insignificant after controlling for total assets

[5]Comparing the differences in wages between large and small firms without controlling for observed and unobserved worker characteristics, Brown and Medoff (1989) and Oi and Idson (1999) find a positive relation of wages to both establishment and firm size.

and various other fixed effects.[6] This indicates that two firms which have the same total assets but different numbers of employees pay the same wage premium. By contrast, the effect of total assets on the wage premium is largely unaffected if we control for total employee fixed effects.

Our third finding is that the wage premium is a concave function of firm size. If we split the overall regression which explains the wage premium by firm size in ten regressions according to the size of the firm, we find that the impact of firm size is particularly strong for smaller firms and declines with firm size. For the largest decile of firms, the impact of firm size on the wage premium is statistically and economically insignificant. It is worth to note that the tenth decile consists of approximately the largest 3,000 German firms (for comparison, about 500 firms are listed at the German stock exchange in Frankfurt). These firms have average total assets of about 260 million Euro, and the smallest firm in the 10th decile has about 100 million Euro total assets. Thus, the overall LFWP is entirely driven by comparatively small firms with less than 100 million total assets. Once a firm reaches a critical size threshold, a further increase in size does not lead to a higher wage premium.[7]

To obtain the fourth stylized fact, we add controls for firm-level rents in the wage premium - size regressions. More precisely, we control for firm profitability (return on assets), employee productivity (sales per employee), and market power of the firm (market share). We either control for those factors by including fixed effects, linear controls, or quadratic controls in the regressions. Both the statistical significance and the magnitude of the firm size coefficient remain unchanged after controlling for these factors. This result provides evidence that firm-level rents cannot explain the existence of the LFWP. Rather, larger firms pay higher wages independently of the rents that they generate.

The fifth stylized fact is related to the time-variation of the wage premium. In our main models, we use data from 2010 to 2016 to estimate the regressions.

[6]In our main models, we include total-assets fixed effects to control for the effect of total assets on the dependent variable. Thus, we are essentially comparing firms with the same level of total assets. Total assets fixed effects are constructed by forming 100 groups (percentiles). This approach has the advantages that it also captures non-linear effects and avoids collinearity issues. The results are similar if we add total assets as independent variable instead of using total assets fixed effects.

[7]This finding is in line with prior evidence by Abowd, Kramarz and Margolis (1999) who document that the firm wage premium of French firms is increasing at a decreasing rate with total employees.

Now we go back until 1993 and exploit variation of the wage premium over time.[8] We find that the coefficient for firm size drops by about one-third if we only exploit time-series variation by adding establishment-fixed effects. This result indicates that part of the LFWP can be explained by time-constant establishment of firm characteristics.

Lastly, we investigate whether reverse causality can explain the LFWP. In this perspective, firms pay higher wage premiums would grow faster, leading to a positive correlation between firm size and the wage premium. To investigate this possibility, we exploit a setting in which firm growth is driven by exogenous factors. In particular, we use variations in the worldwide sales of the German car manufacturers Audi, BMW, Mercedes-Benz, Porsche, and Volkswagen as instruments for the growth of firms which are located close to one of their factories. An increase in sales of these car manufacturers leads to more economic activities at the locations of their factories and thus to a growth of the non-tradable sector there. At the same time, sales of cars on the world market is unlikely to have any direct consequences for the non-tradable sector in regions with car factories.[9]

The next step is to link these stylized facts to the theoretical explanations of the LFWP. Although some of our findings are consistent with rent sharing, others are not. The findings that firm size, not establishment size matters for the wage premium and that assets, not employees drive the premium are consistent with this theory. Firms' productivity may increase as a function of firm size (not establishment) size, e.g., due to economies of scale or market power. Similarly, an increase productivity is likely more closely linked to increases in capital (total assets) than a higher number of employees. However, other results contradict the rent sharing theory. The finding that the wage premium is a concave function in firm size is not consistent with the idea that the most productive firms with very high market power ("superstar firms") pay a wage premium. Furthermore, the effect of firm size on wage remains virtually unchanged if we control for firm profitability, labor productivity, or market power, which provides strong evidence against the rent sharing theory.

[8]For this test, the wage premium is estimated for the intervals 1993 to 1999, 1998 to 2004, 2003 to 2010, and 2010 to 2016. Because we use multiple intervals for this test and wage premiums are not comparable between intervals, the dependent variable is a z-score.

[9]Please note that approximately 80 percent of the cars produced by Audi, BMW, Mercedes-Benz, Porsche, and Volkswagen are sold outside Germany.

By contrast, our results are in line with efficiency wage theories. In particular, they support the idea that firms pay a wage premium to incentives workers who cannot be perfectly monitored. Difficulties in monitoring likely increase in firm size, not establishment. Similarly, monitoring likely becomes more challenging if the firm as a whole grows, even if the number of employees remains constant. Furthermore, the increase in monitoring cost may decrease with firm size. For small firms with few employees monitoring cost increase rapidly if the firm grows. The reason is that small firms do not yet have any mechanisms in place that allow them to monitor employees efficiently (e.g., hierarchy layers, management tools like balanced scorecards, or IT systems). For larger firms, the increase in monitoring cost if the firm grows may be less pronounced because the firm has already professionalized the monitoring of employees. The fifth finding that time-constant omitted variables explain part of the size-wage premium relation is consistent with differences in the monitoring cultures across firms. Overall, our findings are consistent with the idea that firms pay efficiency wages to incentive workers but fail to support rent-sharing explanations.

2. Data & Method

2.1. Employer-employee Data

We use administrative matched employer-employee data from the German social security system 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 include 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.[10] To construct our sample, the data preparation closely mirrors the steps conducted by Card, Heining and Kline (2013) (henceforth, CHK). We first obtain information on the universe of full-time jobs held by male workers with age 20-60 in East and West Germany from

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

1993 to 2016.[11] 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. 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. Since the wage data tracks earnings only up to a certain threshold, the contribution assessment ceiling ("Beitragsbemessungsgrenze"), we follow the procedure suggested by Dustmann, Ludsteck and Schönberg (2009) 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 years, 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 final data set consists of over 180 million worker-establishment-year observations, roughly equally distributed over the four intervals. It comprises over 17 million workers and over 3.8 million establishments.

2.2. Implementation of AKM-type Regression

To isolate the wage premium paid by an establishment, we follow the two-way fixed effects wage regression, pioneered by AKM. The regression model that we estimate is

$$y_{i,t} = \alpha_i + \psi_{J(i,t)} + \beta X_{i,t} + r_{i,t}, \tag{1}$$

where $y_{i,t}$ denotes the log average daily wage of individual i in year t, α_i the person fixed effect, $\psi_{J(i,t)}$ an establishment fixed effect, $x'_{i,t}$ an index of time-varying observable characteristics and $r_{i,t}$ an error component. $X_{i,t}$ includes an unrestricted set of year dummies as well as quadratic and cubic terms in age fully interacted with educational attainment. The function $J_{i,t}$ gives the identity of the unique establishment that employs worker i in year t.

We estimate equation 1 separately for the four overlapping sample intervals: 1993-1999, 1998-2004, 2003-2010, 2010-2016. For each of the intervals, the estimation is done on the largest connected set of establishments based on worker flows between the establishments. From our estimations, we obtain a series of $\psi_{j,t'}$, establishment fixed effect, for each establishment j over the four

[11] The IEB data do not include employment spells of civil servants and self-employed workers.

intervals t'. We interpret the establishment fixed effect as the wage premium an establishment j pays to all employees in interval t'.

2.3. Firm data

The matched employer-employee data provides only information on employees and establishments. E.g., individual i is employed at establishment j. But, it does not include information on the firm structure. Therefore, we use a novel data set ORBIS-ADIAB 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 a single firm k.

This linkage process 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 establishment / company name as well as legal form are the most important ones. After extensive pre-processing, record linkage techniques are applied. The linkage of establishment to firm identifiers is carried out for the years 2006 to 2013. In this period, at least one establishment could be assigned to 975,880 of the 1,948,788 Orbis companies which equals a success rate of about 50%.[12]

To build up a long-time matched firm-establishment data set, we must make one additional assumption. We use this linking table between establishment and firm identifiers that is developed on data from 2006 to 2013 for our entire sample period from 1993 to 2016. Hence, we assume that an establishment that belongs to a firm between 2006 to 2013 also belongs to this firm in the 1993-2016 period. For establishments that are not involved in m&a activities or ownership changes, we expect this assumption to be valid. But even for establishments undergoing ownership changes, the number of false positives should be limited since such activities are likely to result in a new establishment identifiers.[13]

[12] The novelty of this data set also comes with some limitations. The information form Orbis used in the linking process is extracted on January 30, 2014. This means that time-static information from Orbis is used for the matching over time from 2006 to 2013. E.g., it is constantly used the latest company name (from 2014) for the entire period from 2006 to 2013 and not the company from the respective years.

[13] If the ownership change happens before 2006, only the latest establishment identifier will be used in the linkage process. Hence, the old establishment will not be covered in our data set. If the ownership change happens between 2006 and 2013, both establishment

Firm-level financial information from 1993 to 2016 is derived from the Orbis database by BvD. We exclude firms from our sample for which we do not obtain information on total assets and total sales. As we observe the establishment wage premia on interval level, we must also aggregate the firm-level financial information. To do this, we calculate means over the respective intervals. Using the BvD identifier, we link the financial information from Orbis to the data on the establishment wage premium. Our final data set is on the firm-establishment-interval level.

2.4. Analyzing the Wage Premium

Our main goal is to examine why the large firm wage premium exists. Technically, the establishment wage premium from the AKM-type regression is estimated as an establishment fixed effect and can only be interpreted relative to an omitted establishment. Our main sample is the cross-section of the latest interval (2010-2016).[14] Hence within this interval, all establishment wage premia are interpreted relative to the same omitted establishment. Since we observe the wage premium on establishment level, we can control for unobserved geographic and industry effects (e.g., cost of living) within a firm. For the cross-sectional analysis, our standard regression specification is,

$$\psi_{j,k} = \beta z_{j,k} + \pi_j + \rho_k + \tau_j + r_j, \tag{2}$$

where $\psi_{j,k}$ denotes the establishment wage premium of establishment j that belongs to firm k, $z_{j,k}$ a size indicator on establishment- or firm-level, πj establishment-industry fixed effects, ρ_k firm-industry fixed effects and τ_j establishment-district fixed effects.

As size indicators, we use the establishment's number of employees, the firm's total number of employees, and the firm's total assets. If we want to compare the relevance of these size measures for the existence of the LFWP, we

identifiers (old and new) are used in the linkage process. Only for Establishment involved in ownership changes occurring from 2014 on, we may falsely assume in the years 2014-2016 that an establishment still belongs to the former owner. For details on the establishment identifiers and possible changes of establishment identifiers over time, please refer to the technical report by Schmucker et al. (2016).

[14]We choose the sample interval 2010-2016 since the availability of firm-level financial information substantially increases over our sample time and this is then the interval with the largest number of firms. Nevertheless, we obtain similar results if we reproduce our analysis on the other intervals.

face the issue that the size indicators are highly correlated. Figure 1 presents the visualization of a correlation matrix. E.g., total employees and total assets show a correlation of 0.80. To avoid collinearity issues, we do not include both total employees and total assets as control variables into one regression model. Instead, we include only one of the two size measures as independent variable in the regression and control for the other size measure by 100 group fixed effects (percentiles). E.g., we test whether total assets is still able to explain the wage premium within a group of establishments that belong to firms with a comparable number of total employees.

In an additional analysis, we explore the within-establishment variation of the wage premium over the four sample intervals. To make the establishment wage premia of the four intervals comparable, we calculate the z-score of the establishment wage premium for each interval.[15] This transforms the wage premium into a variable with mean equal to zero and a standard deviation of one. For this analysis, our regression specification is,

$$zscore(\psi_{j,k,t'}) = \beta z_{j,k,t'} + \tau_{t'} + \phi_j + r_j, \tag{3}$$

where $zscore(\psi_{j,k,t'})$ denotes the z-score of the establishment wage premium in interval t', $z_{j,k,t'}$ a size indicator on establishment- or firm-level in interval t', $\tau_{t'}$ interval fixed effects, and ϕ_j establishment fixed effects.

In a final analysis, we deal with the possibility of reverse causality as explanation for the LFWP using an instrument variable approach. The idea is to exploit the growth of the German car manufacturers that primarily comes from the world market as instrument for the growth of local firms. Our concrete instrument is the number of delivered vehicles by Audi, BMW, Mercedes-Benz, Porsche and Volkswagen. These manufacturers run 27 factories with more than 1000 employees in 25 districts of Germany (see Appendix B.1 for a complete list of the car factories).

In this context, the exclusion restriction requires that local firms are affected from the growth of German car manufacturers only through the local demand of the car factories and their employees. Therefore, we define "local

[15]In detail, we subtract the mean wage premium of the interval from the establishment wage premia and divide this difference by the interval's standard deviation of the wage premium.

firms" as firms from non-tradable industries[16], hotels and similar accommodations, plus firms providing consultancy and support services that might be of relevance for the local car factory but not for the world market (see Appendix B.2 for a complete list of the industries).[17] To be considered in the IV analysis, local firms must be headquartered in the same district as a car factory. For local firms that satisfy this condition, we include all establishments in these districts.

2.5. Descriptive Statistics

Table 1 provides descriptive statistics for the establishments in the latest sample interval from 2010-2016. In this interval, we observe 218,260 establishments that belong to 169,173 firms. The wage premium and all financial variables are winsorized at the 1% level.[18] The median establishment is an one-establishment firm. The average establishment has about 11 employees and belongs to a firm with 25 employees and total assets of about EUR 3m. To rule out any time effects, we always use cpi-adjusted total assets in m of 2013 EUR. A detailed description of all variables can be found in Appendix A.1.

3. Results

3.1. Firm Size and the Wage Premium

Figure 2 illustrates the relationship between wage premium and size. For size, we use the establishment-level number of employees, firm total employees and firm total assets to construct ten size groups (deciles). The wage premium is interpreted relative to the mean of the sample interval 2010-2016.

In Subfigure (a), the deciles are based on the establishment number of employees. The graph indicates a linear relationship between wage premium and establishment size with the strongest increase of the wage premium from

[16]The starting point is the definition of non-tradable industries by Mian and Sufi (2014). We translate the 4-digit NAICS codes to the WZ 2008 industry definition. Given our instrument from the automotive industry, we do not include retail of motor vehicles.

[17]We think that this instrument fits well to our setting of four time intervals of 7 to 8 years. Over these large time spans, the local demand stemming from car factories should be an important determinant for the growth of local firms.

[18]We winsorize the financial variables from Orbis in the firm-year data set at the 1% before aggregating the financial variables by the mean on interval-level.

the second largest to the largest establishments. While the smallest establishments pay a 5% lower wage premium, the wage premium of the largest establishments is about 10% above the mean. In Subfigure (b), we use the fim's total employees as size measure. The graph looks very similar to Subfigure (a). In Subfigure (c), the size deciles are constructed by firm total assets. Again, we observe a pretty linear relationship but with a greater slope. The establishments belonging to the smallest firms pay a more than 10% lower wage premium and the establishments of the largest firms a 10% larger wage premium than the sample mean.[19]

3.2. Is it a Large Establishment or Large Firm Wage Premium?

In this section, we test whether it is a large establishment or a large firm wage premium. Therefore, we regress the establishment wage premium separately on the establishment's number of employees, the firm's total employees and the firm's total assets. Table 2 presents the results.

In Panel A, we use number of employees as establishment size measure. Column 1 presents the outcome of an OLS regression. The regression coefficient on establishment size is 0.029. Hence, if the establishment size doubles, the wage premium increases by 2.9%. Next, we introduce firm fixed effects to explore whether an establishment size premium still exists within one firm. In specifications with firm fixed effects, our sample consists only of the 62,242 establishments that belong to multi-establishment firms. The coefficient substantially drops to 0.006. If we further add establishment-industry and -district fixed effects (Column 5), the coefficient reduces to 0.003. Now, an increase of establishment size by 100% leads only to a rise of the wage premium by 0.3%. From this result, we conclude that there is no economically meaningful establishment size wage premium.

In Panel B, we use total employees as firm size measure. The coefficient of the OLS regression is 0.018. In the following columns, we introduce 100 establishment-size fixed effects (percentiles) to test whether a LFWP exists for establishments of similar size. In Column 5, we further introduce fixed effects

[19]Recent papers analyze the development of the difference in average earnings between large and small firms over time for different countries. Bloom et al. (2018) documents that the average pay gap substantially declined for the U.S. Colonnelli et al. (2018) look at Brazil, Germany, Sweden and UK. For Germany, they they do not find a decrease in the average pay gap. Consistent with this, we do not observe a decline of the Large Firm Wage Premium (net of employee characteristics) over our sample period.

for the industry of the firm, the establishment industry and the establishment district. The coefficient increases to 0.024. In Panel C, we run the same analysis with total assets as firm size measure. For all models, the coefficient ranges between 0.029 and 0.030. These results provide evidence for a large firm wage premium.

3.3. How to Measure Firm Size? Total Employees vs. Total Assets

Next, we analyze whether total employees or total assets seems to be the more relevant firm size measure in the context of the LFWP. Since total employees and total assets show a correlation of 0.80, we must be careful to rule out potential collinearity issues. This is why, we do not include both variables into one regression model. Instead, we include only one of the two variables and use the other size measure to form 100 size groups (percentiles) that we use as fixed effects.

Table 3 presents the results. In Panel A, we use the firm's total employees as size measure. Column 1 repeats the plain OLS analysis which leads to a coefficient of 0.018. In Columns 2, we introduce the 100 size-group fixed effects based on total assets, the coefficient on total employees becomes negative with -0.017. If we further add firm-industy, establishment-industry and establishment-district fixed effects (Column 5), the coefficient is still negative but pretty close to zero with -0.003. Hence, if we control for total assets, there is no positive relationship between the wage premium and total employees.

In Panel B, the size measure is total assets and we control for total employees with 100 size group fixed effects. Column 1 presents the OLS coefficient that is 0.029. The addition of the 100 size-group fixed effects controlling for total employees even increases the coefficient to 0.046 (Column 2). In the final model with all fixed effects (Column 5), the coefficient is 0.036. This can be interpreted as follows: If a firm keeps the number of employees more or less constant (remains in the same size group) but increases total assets by 100%, the wage premium increases by 3.6%. Thus, total assets as a proxy of capital seems to be the driving force behind the LFWP.

3.4. Is the Large Firm Wage Premium Linear in Firm Size?

Subfigure (c) of Figure 2 indicates a linear relationship between the wage premium and a firm's total assets. In this section, we test this in a more formal way. Therefore, we split the latest interval 2010-2016 into deciles based on total assets. We separately run our standard regression model (Eq. 2) for each of the deciles.

Table 4 presents the results, for the smallest deciles (1st to 5th) in Panel A and for the largest deciles (6th to 10th) in Panel B. The coefficients range from 0.048 (2nd decile) to 0.009 (10th decile). Overall, we observe a pretty clear pattern that the coefficient on total assets is decreasing over the deciles. In other words, the LFWP seems to be concave.

The concave shape of LFWP is visualized in Figure 3. The y-axis shows the wage premium relative to the sample mean. The x-axis presents average firm size of the 10 firm size groups. In Subfigure (a) the firm-size measure is ln(total assets) as in Table 4. Here, you can clearly see that the increase of the wage premium with total assets is most pronounced for small firms and decreases with the size of the firms. The concave shape becomes even stronger, if we use the non-logarithmic values of total assets. From the 1st to the 9th decile, average total assets increases from EUR 0.20m to about 39m. However, the jump from the 9th to the 10th deciles is substantially. In the 10th decile, the average total assets is about 260m.[20]

3.5. Can Firm-level Rent Measures Explain the Large Firm Wage Premium?

A main benefit of our matched firm-establishment data set is that we can measure rent on the firm level. We use three proxies for rent. These are return on assets as a measure of operating profitability, total sales over employees as (accounting-based) productivity measure, and market share as a measure of a firm's market power. To examine whether the magnitude of the rent can explain the LFWP, we add these measures as control variables to our regression models. Table 5 presents the results. For each of the variables, we follow three approaches to control for it. In Column 1, we form 100 groups (percentiles) based on the rent measure and include these as control-group fixed effects. In Column 2 and 3, we add the rent measure as control variable, and as plain plus quadratic control variable.

In Panel A, the control variable is return on assets. The addition of return on assets has a minor effect on the coefficient on total assets. It still ranges from 0.030 to 0.032. In our baseline model without control variable, the coefficient is

[20]This is still small if compared to largest German firms. The median DAX30 company, for instance, has total assets of approx. 40,000m. The corresponding value for the smallest DAX30 company is approx. 3,000m.

0.030 (see Column 5 of 2). Panel B presents the results for sales per employee. Again, the coefficients with 0.031 remain in close range to our baseline model. In Panel C, we use the firm's market share. In Column 1, we include 100 control-group fixed effects based on the market share. The coefficient decreases to 0.026. If we include market share as linear and quadratic control variable, the coefficient increases to 0.032. We find that none of the three rent measures has a substantial effect on the coefficient on total assets. These results make it unlikely that higher rents of large firms are the main explanation for the existence of the LFWP.

3.6. To what extent is the Large Firm Wage Premium driven by time-constant omitted variabes?

In this section, we examine whether (omitted) time-invariant characteristics can explain the LFWP. We exploit the within-establishment variation of the wage premium across our four sample intervals in order to explain which degree of the LFWP is related to time-invariant unobserved (firm) characteristics.

Table 6 presents the results. The dependent variable is the z-score of the wage premium to make the wage premium comparable across the four time intervals (for further details see Section 2.4). Column 1 presents the outcome of a pooled OLS regression. The coefficient on total assets is 0.118. This coefficient can be interpreted as follows: If the firm size increases by 100%, the wage premium rises by 0.118 standard deviations. In Column 2, we add time-interval fixed effects. The coefficient increases to 0.136. These coefficients are in line with our previous OLS regressions on the latest sample interval (2010-2016). There, the OLS coefficient is 0.029 which equals 0.142 standard deviations.[21].

In Column 3, we add establishment fixed effects. Now, the identification exclusively comes form the within-establishment variation over time. Compared to the OLS regressions, this leads to a decrease of the coefficient on total assets to 0.072. From this, we conclude that about one third of the LFWP can be explained by time-invariant unobserved firm characteristics.

[21]The standard deviation of the wage premium is 0.204 in the interval 2010-2016 (see Table 1)

3.7. Is the LFWP caused by Reverse Causality?

In this section, we apply an instrumental variable approach to identify the LFWP. We use the growth of the German car manufacturers that primarily comes from the world market as instrument for the growth of local firms. Our concrete instrument is the number of delivered vehicles by Audi, BMW, Mercedes-Benz, Porsche and Volkswagen. We define local firms as firms operating in non-tradable industries, hotels and similar accommodations, plus providers of consultancy and support services that are located in the same district as a car factory of these manufacturers (see Section 2.4 for a discussion of the IV approach, Appendix B.1 for a list of the car factories, and Appendix B.2 for a list of the used industry codes).

Table 7 presents the results. In Column 1, we run a plain OLS regression for the sample of local firms that we use in the subsequent 2SLS regressions. The coefficient on total assets is 0.202. Compared to the full sample (0.118, Column I of Table 6), the magnitude of the effect is larger for local firms. If a local firm doubles in total assets, its wage premium increases by 0.202 standard deviations.

We run two IV specifications. In the first one (Columns 2a/2b), we use all local firms. In the second one (Columns 3a/3b), we consider only districts in which the automotive industry accounts for 5% of the overall employment and local firms that employ at least 50% of their workers in this district. Delivered cars have a strong economic and statistical significance in explaining the growth of local firms' total assets (Columns 2a/3a). On the second stage (Columns 2b/3b), we observe again a positive relation between (instrumented) total assets and the wage premium. These coefficients with 0.804 and 0.884 are larger than the pooled OLS model. This setting should mitigate the concern that the LFWP results from reverse causality.

4. Conclusion

Despite ample evidence about higher wages in larger firms, little is know about the reasons for this result. Using a matching of administrative employee-level information from Germany with firm-level data allows us to analyze how size affects the wage premium after adjusting for worker characteristics. The final dataset covers 218,260 establishments that belong to 169,173 firms in the period 2010 to 2016.

We start by documenting six stylized facts about the wage premium in large firms. We find that (i) firm size, not establishment size matters for the LFWP, (ii) the LFWP is driven by assets, not employees, (iii) the premium is a concave function of firm size, (iv) firm profitability, market share, and productivity have little impact on the premium's magnitude, (v) time-constant omitted variables explain about one-third of the impact of size on the premium, and (vi) reverse causality is unlikely to cause the LFWP.

The next step is to link these stylized facts to the theoretical explanations of the LFWP. Although some of our findings are consistent with rent sharing, others are in stark contrast. In particular, we find that the effect of firm size on the wage premium is unaffected by controls for firm profitability, market power, or productivity. Thus, our findings do not support rent sharing explanations of the LFWP. Rather, our results are in line with efficiency wage theories. In particular, they support the idea that firms pay a wage premium to incentivize workers who cannot be perfectly monitored.

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.

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., 2017. The Fall of the Labor Share and the Rise of Superstar Firms. NBER Working Paper.

Bebchuk, L., Cremers, M., Peyer, U., 2011. The CEO pay slice. Journal ofFinancial Economics 102 (1), 199–221.

Bebchuk, L., Fried, J., 2004. Pay without performance: The unfulfilled promise of executive compensation. Cambridge, MA: Harvard University Press.

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.

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

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.

Colonnelli, E., Tåg, J., Webb, M., Wolter, S., 2018. A Cross-Country Comparison of Dynamics in the Large Firm Wage Premium. AEA Papers and Proceedings 108, 323–27.

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.

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

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

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.

Gabaix, X., Landier, A., 2008. Why has CEO Pay Increased So Much? Quarterly Journal of Economics 123 (1), 49–100.

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.

Mian, A., Sufi, A., dec 2014. What Explains the 2007–2009 Drop in Employment? Econometrica 82 (6), 2197–2223.

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

Ouimet, P., Zarutskie, R., 2014. Who works for startups? The relation between firm age, employee age, and growth. Journal of Financial Economics 112 (3), 386–407.

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

Schmucker, A., Seth, S., Ludsteck, J., Eberle, J., Ganzer, A., Aug. 2016. Establishment History Panel 1975-2014. FDZ Datenreport. Documentation on Labour Market Data 201603_en, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Song, J., Price, D. J., Guvenen, F., Bloom, N., von Wachter, T., 2018. Firming up Inequality. Quarterly Journal of Economics, forthcoming.

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

Figure 1
Correlation matrix
This figure presents the visualization of the correlation matrix. The sample is the latest interval 2010-2016. A detailed description of all variables can be found in Appendix A.1.

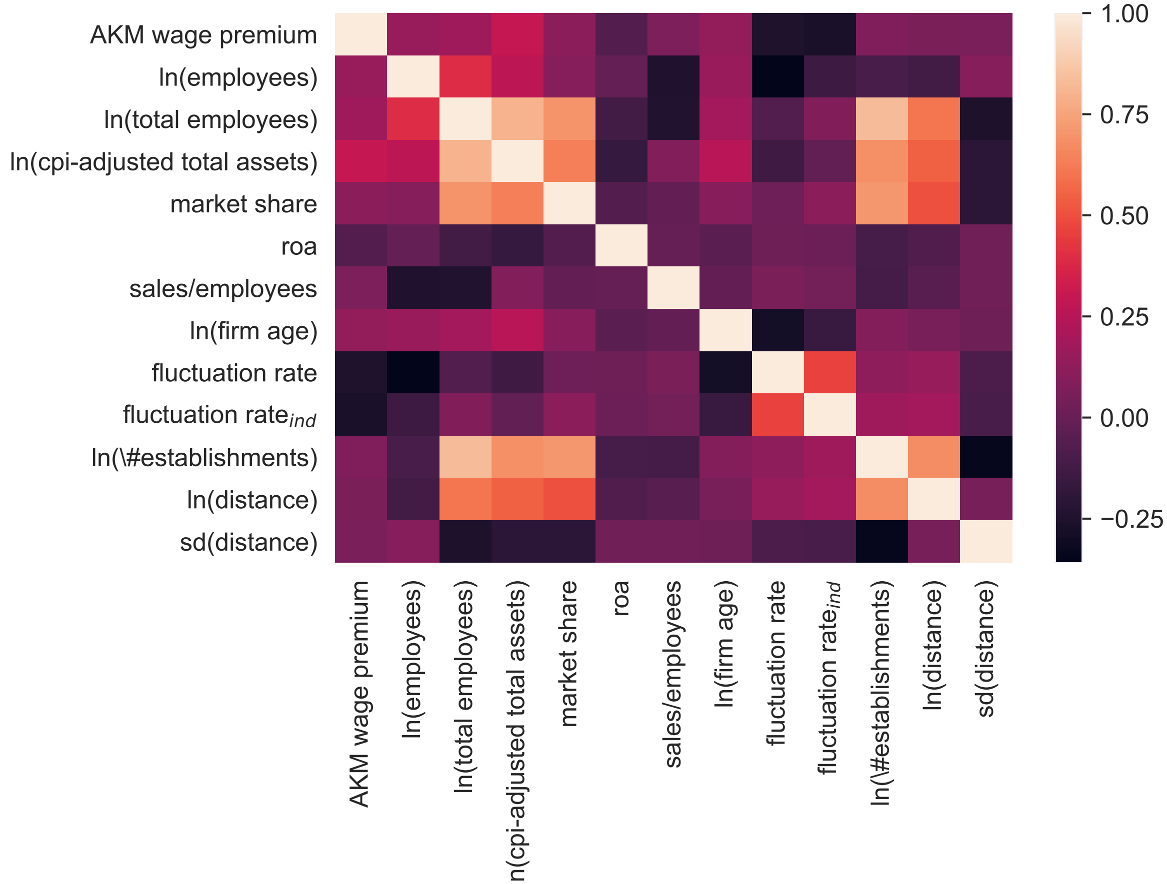
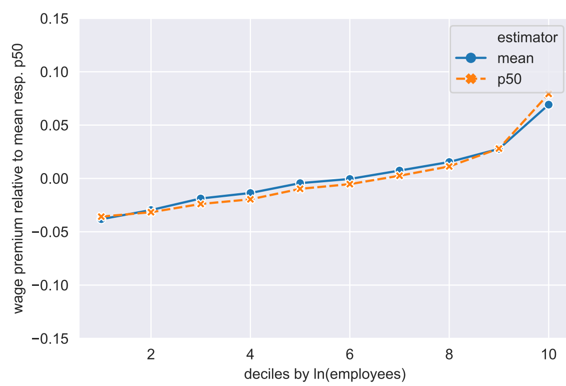


Figure 2

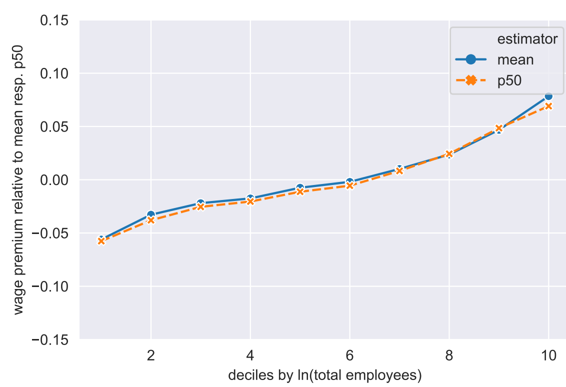
Wage premium by establishment and firm Size

This figure presents the establishment wage premium by deciles of establishment and firm size. The y-axis shows the wage premium relative to the sample mean. The x-axis shows the deciles (1 to 10) based on establishment or firm size. The deciles in subfigure (a) are constructed by the establishment number of employees. The deciles in subfigure (b) use total employees of the firm and the deciles in subfigure (c) are based on total assets of the firm. The sample is the latest interval 2010-2016. A detailed description of all variables can be found in Appendix A.1.

(a) deciles by employees



(b) deciles by total employees



(c) deciles by total assets

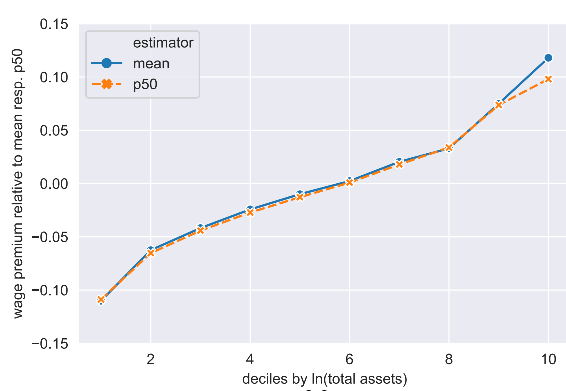
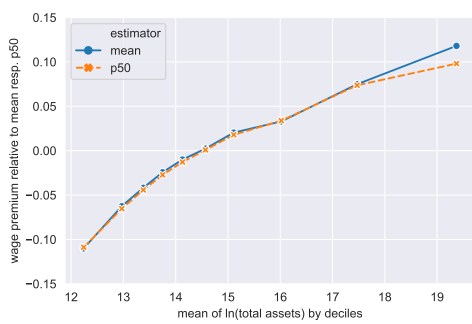
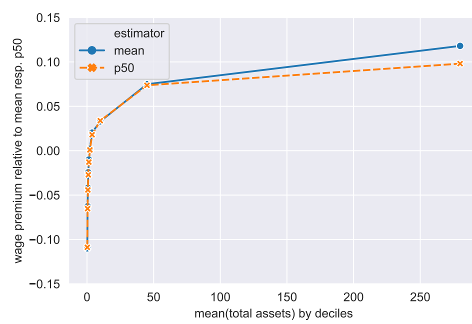


Figure 3
Shape of the LFWP
This figure presents the establishment wage premium by deciles and firm size. The y-axis shows the wage premium relative to the sample mean. The x-axis shows the mean firm size of the firm size deciles. The firm size measures are ln(total assets) in Subfigure (a), total assets in Subfigure (b), ln(total employees) in Subfigure (c), and total employees in Subfigure (d). The sample is the latest interval 2010-2016. A detailed description of all variables can be found in Appendix A.1.

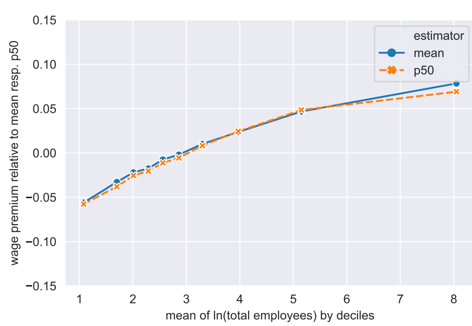
(a) ln(total assets)



(b) total assets



(c) ln(total employees)



(d) total employees

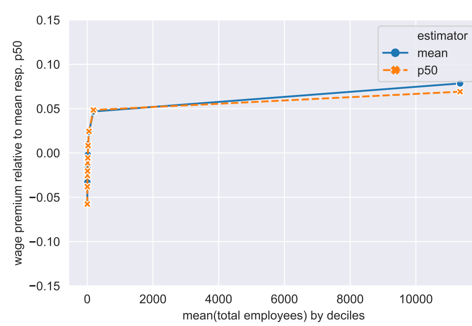


Table 1
Descriptive statistics

This table presents descriptive statistics of the establishments in the latest sample interval from 2010 to 2016. Reported are the number of observations (Obs), mean value (Mean), standard deviation (SD) 25% percentile (p25), median (p50), 75% percentile (p75). A detailed description of all variables can be found in Appendix A.1.

	Obs	Mean	SD	P25	P50	P75
AKM wage premium	218260	0.276	0.204	0.158	0.295	0.409
ln(employees)	218260	2.363	1.137	1.609	2.197	2.996
ln(total employees)	218260	3.233	2.052	1.946	2.639	3.932
ln(cpi-adjusted total assets)	218260	14.902	2.095	13.379	14.338	15.991
market share	218260	0.004	0.010	0.000	0.000	0.002
roa	104372	0.123	0.149	0.043	0.102	0.184
sales/employees	218260	0.573	1.020	0.144	0.263	0.543
ln(firm age)	217889	2.809	0.890	2.250	2.862	3.365
fluctuation rate	206100	0.384	0.249	0.213	0.313	0.478
fluctuation rate$_{ind}$	206100	0.388	0.119	0.314	0.370	0.430
ln(#establishments)	218260	1.315	1.452	0.693	0.693	1.099
ln(distance)	217914	1.142	2.124	0.000	0.000	0.000
sd(distance)	62115	1.810	1.104	1.151	1.615	2.452

Table 2
Is the LFWP driven by establishment or firm size?

The dependent variable is the establishment wage premium. In Panel A, we use the establishment's number of employees as size measure. In Panel B, the size measure is the firm's total employees. In Panel C, the firm's total assets is the size measure. The sample is the latest interval 2010-2016. 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: large establishment wage premium					
ln(employees)	0.029***	0.006***	0.005***	0.004***	0.003***
	(20.002)	(6.661)	(6.746)	(3.383)	(2.809)
Firm FE	No	Yes	Yes	Yes	Yes
Estab industry FE	No	No	Yes	No	Yes
Estab District FE	No	No	No	Yes	Yes
Obs	218,260	62,242	62,029	62,035	62,029
R2	0.026	0.674	0.683	0.691	0.699
Panel B: large firm wage premium - total employees					
ln(total employees)	0.018***	0.014***	0.021***	0.015***	0.024***
	(8.083)	(8.123)	(18.777)	(8.424)	(23.619)
Estab size FE	No	Yes	Yes	Yes	Yes
Firm industry FE	No	No	Yes	No	Yes
Estab industry FE	No	No	Yes	No	Yes
Estab District FE	No	No	No	Yes	Yes
Obs	218,260	218,260	218,007	218,043	218,007
R2	0.032	0.044	0.226	0.162	0.331
Panel C: large firm wage premium - total assets					
ln(total assets)	0.029***	0.029***	0.030***	0.029***	0.030***
	(18.308)	(18.193)	(31.839)	(19.226)	(35.382)
Estab size FE	No	Yes	Yes	Yes	Yes
Firm industry FE	No	No	Yes	No	Yes
Estab industry FE	No	No	Yes	No	Yes
Estab District FE	No	No	No	Yes	Yes
Obs	218,260	218,260	218,007	218,043	218,007
R2	0.092	0.103	0.258	0.221	0.360

Table 3
Is the LFWP driven by employees or assets?

The dependent variable is the establishment wage premium. In Panel A, we use the firm's total employees as size measure and introduce 100 size groups (percentiles) based on total assets. In Panel B, the size measure is the firm's total asssets and we introduce 100 size groups (percentiles) based on the total employees. The sample is the latest interval 2010-2016. 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: total employees					
ln(total employees)	0.018***	-0.017***	-0.009***	-0.013***	-0.003***
	(8.083)	(-9.749)	(-9.177)	(-7.233)	(-3.417)
Total assets FE	No	Yes	Yes	Yes	Yes
Firm industry FE	No	No	Yes	No	Yes
Estab industry FE	No	No	Yes	No	Yes
Estab District FE	No	No	No	Yes	Yes
Obs	218,260	218,260	218,007	218,043	218,007
R2	0.032	0.109	0.258	0.220	0.359
Panel B: total assets					
ln(total assets)	0.029***	0.046***	0.039***	0.045***	0.036***
	(18.308)	(60.099)	(60.663)	(60.637)	(58.145)
Total employees FE	No	Yes	Yes	Yes	Yes
Firm industry FE	No	No	Yes	No	Yes
Estab industry FE	No	No	Yes	No	Yes
Estab District FE	No	No	No	Yes	Yes
Obs	218,260	218,260	218,007	218,043	218,007
R2	0.092	0.115	0.260	0.230	0.361

Table 4
Is the LFWP linear in firm size?

The dependent variable is the establishment wage premium. We split the sample into deciles based on total assets. For each decile, we run Eq. 2 with ln(total assets) as firm size measures. Panel A presents the 1st to 5th deciles (smallest firms) and Panel B the 6th to 10th deciles (largest firms). The sample is the latest interval 2010-2016. 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: 1st to 5th deciles by total assets					
ln(total assets)	0.040***	0.048***	0.038***	0.031***	0.037***
	(10.493)	(4.985)	(3.425)	(2.941)	(3.679)
Firm industry FE	Yes	Yes	Yes	Yes	Yes
Estab industry FE	Yes	Yes	Yes	Yes	Yes
Estab District FE	Yes	Yes	Yes	Yes	Yes
Obs	21,741	21,759	21,765	21,777	21,763
R2	0.261	0.277	0.317	0.321	0.332
Panel B: 6th to 10th deciles by total assets					
ln(total assets)	0.025***	0.042***	0.026***	0.018***	0.009
	(3.110)	(7.138)	(6.311)	(4.092)	(0.647)
Firm industry FE	Yes	Yes	Yes	Yes	Yes
Estab industry FE	Yes	Yes	Yes	Yes	Yes
Estab District FE	Yes	Yes	Yes	Yes	Yes
Obs	21,768	21,778	21,767	21,770	21,720
R2	0.362	0.382	0.414	0.412	0.474

Table 5
Can firm-level rent measures explain the LFWP?

The dependent variable is the establishment wage premium. Panel A controls for return on assets, Panel B for sales/employees, and Panel C for market share. In each specification, we run Eq. 2 with ln(total assets) as firm size measure plus three types of controls for the respective control variable. Column 1 adds 100 group fixed effects, Column 2 the control variable, and Column 3 the plain control variable and the quadratic term. The sample is the latest interval 2010-2016. 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)
Panel A: return on assets			
ln(total assets)	0.032***	0.030***	0.031***
	(47.505)	(36.661)	(35.760)
return on assets		-0.013	-0.068***
		(-1.559)	(-7.827)
return on assets^2			0.159***
			(8.385)
Control variable FE	roa	-	-
Firm industry FE	Yes	Yes	Yes
Estab industry FE	Yes	Yes	Yes
Estab District FE	Yes	Yes	Yes
Obs	104,200	104,200	104,200
R2	0.418	0.412	0.413
Panel B: sales/employees			
ln(total assets)	0.031***	0.031***	0.031***
	(47.263)	(39.842)	(39.972)
sales/employee		0.006***	0.010***
		(9.617)	(5.815)
sales/employee^2			-0.001***
			(-2.896)
Control variable FE	sales/empl	-	-
Firm industry FE	Yes	Yes	Yes
Estab industry FE	Yes	Yes	Yes
Estab District FE	Yes	Yes	Yes
Obs	218,007	218,007	218,007
R2	0.359	0.357	0.357

continued on next page

Table 5 continued

Panel C: market share			
ln(total assets)	0.026***	0.033***	0.033***
	(31.338)	(52.998)	(46.107)
market share		-0.838***	-0.094
		(-3.094)	(-0.103)
market share^2			-18.237
			(-0.744)
Control variable FE	market share	-	-
Firm industry FE	Yes	Yes	Yes
Estab industry FE	Yes	Yes	Yes
Estab District FE	Yes	Yes	Yes
Obs	218,007	218,007	218,007
R2	0.360	0.357	0.357

Table 6
Is the LFWP caused by time-constant omitted variables?

The dependent variable is the z-score of the establishment wage premium. The z-score is separately constructed for each interval. It transforms the wage premium into a variable with mean equal to zero and a standard deviation equal to one. For further details, see Section 2.4. The sample consists of establishments from all four intervals from 1993 to 2016 that we observe for at least two intervals. 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)
ln(total assets)	0.118***	0.136***	0.072***
	(10.764)	(14.825)	(3.621)
Interval FE	No	Yes	Yes
Establishment FE	No	No	Yes
Obs	257,556	257,556	257,556
R2	0.104	0.125	0.832

Table 7
Is the LFWP caused by reverse causality? An IV approach

This table shows the 2SLS instrumental variables regressions with delivered vehicles by Audi, BMW, Mercedes-Benz, Porsche and Volkswagen as instruments for the growth of local firms. Local firms are defined as all firms from non-tradable industries, hotels and similar accomodations, plus providers of consultancy and support services that have their headquarter in the same district as one of the 27 factories of the car manufactures (see Appendix B.2 for a list of the used industries and Appendix B.1 for a list of the car factories). For these firms, we use all establishment-level observations in these distrcits over the four sample intervals from 1993 to 2016. The dependent variable of Column 1, 2b, 3b (2a, 3a) is the z-score of the establishment wage premium (ln(total assets)). Column 1 presents the plain OLS regression on this sample. Columns 2a and 2b present the first and second stage of 2SLS regression using all observations. Columns 3a and 3b present the first and second stage of the 2SLS regression, if we consider only local firms that have 50% of their total employees in the district and districts for which the automotive industry accounts for at least 5% of total employement. 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.

Model	(1) OLS	(2a) IV 1st	(2b) IV 2nd	(3a) IV 1st	(3b) IV 2nd
ln(total assets)	0.202***		0.802***		0.884***
	(21.867)		(5.053)		(3.349)
ln(vehicles)		0.384***		0.274***	
		(6.097)		(3.904)	
Interval FE	No	Yes	Yes	Yes	Yes
Establishment FE	No	Yes	Yes	Yes	Yes
Obs	5,650	5,323	5,323	2,095	2,095
K-P rk Wald F statistic		37.17		15.24	

Appendix A. Definition of variables

Table A.1
Definition of Variables

Variable	Description
Main variables	
wage premium	The establishment fixed effect from the AKM-type regression. The implementation and interpretation of the AKM-type regression is explained in detail in Section 2.2. Source: IAB, own estimation.
ln(employees)	Natural logarithm of the establishment's full-time male employees. Source: IAB.
ln(total employees)	Natural logarithm of the firm's total number of full-time male employees. Source: IAB.
ln(cpi-adjusted total assets)	Natural logarithm of total assets (toas) in m of 2013 EUR. Source: Orbis.
Other control variables	
roa	Firm's earnings before interest, taxes, depreciation, and amortization standardized by total assets ($\frac{ebta}{toas}$). Source: Orbis.
sales/employees	Firm's total sales divided by the total number of full-time male employees ($\frac{turn}{totalemployees}$). Source: IAB, Orbis.
market share	Firm's total sales standardized by the sum of total sales over all firms operating in a 3-digit WZ 2008 industry ($\frac{sales_k}{\sum_{k=1}^{n} sales_k}$). Source: Orbis.
ln(firm age)	Natural logarithm of a firm's age. Firm age is calculated as the fiscal year - the founding year plus one. Source: Orbis.
fluctuation rate	Establishment's hiring rate plus the separation rate - absolute value of the change in total employees ($abs(HR_t) - abs(SR_t) - abs(\frac{employees_{t+1}}{employees_t} - 1)$. The hiring (separation) rate is defined as the number of employees that flow into (out of) the establishment standardized by the establishment's lagged number of employees ($\frac{inflow_t}{employees_{t-1}}$ / $\frac{outflow_t}{employees_{t-1}}$). Source: IAB.
ind fluctuation rate	The average fluctuation rate in a 3-digit WZ 2008 industry excluding the respective establishment's fluctuation rate. Source: IAB.
ln(distance)	Natural logarithm of the geographic distance in km between the district of the establishment and the district of the firm headquarter. Source: Own estimation.
sd(distance)	Standard deviation of the geographic distance in km between the district of the establishments and the district of the firm headquarter. Source: Own estimation
ln(establishments)	Natural logarithm of the firm's number of establishments. Source: IAB

IAB stands for data provided by the Institute of Employment Research, and Orbis for the Orbis database by Bureau van Dijk.

Appendix B. Details about the IV approach

Table B.1
German Car Factories

This table lists the German car factories by Audi, BMW, Mercedes-Benz, Porsche and Volkswagen. Source: Annual Repors and Company Websites.

plant	brand	employees	founded
BMW Group Werk Dingolfing	BMW	18500	1967
BMW Group Werk Landshut	BMW	4100	1967
BMW Group Werk Leipzig	BMW	5200	2005
BMW Group Werk Munchen	BMW	9000	1917
BMW Group Werk Regensburg	BMW	9000	1984
BMW Group Werk Wackersdorf	BMW	3000	1989
Mercedes-Benz Werk Rastatt	Mercedes-Benz	6500	1992
Mercedes-Benz Werk Sindelfingen	Mercedes-Benz	25000	1915
Mercedes-Benz Werk Bremen	Mercedes-Benz	12500	1978
Mercedes-Benz Werk Berlin	Mercedes-Benz	2500	1902
Mercedes-Benz Werk Hamburg	Mercedes-Benz	2500	1935
Mercedes-Benz Werk Kölleda	Mercedes-Benz	1400	2002
Mercedes-Benz Werk Untertürkheim	Mercedes-Benz	19000	1904
Mercedes AMG GmbH Affalterbach	Mercedes-Benz	1500	1976
Audi Ingolstadt	Audi	44217	1949
Audi Neckarsulm	Audi	16995	1873
Volkswagen Wolfsburg	VW	62268	1938
Volkswagen Emden	VW	9133	1964
Volkswagen Kassel	VW	17027	1957
Volkswagen Hannover	VW	14765	1956
Porsche Werk Zuffenhausen	Porsche	17549	1937
Porsche Werk Leipzig	Porsche	4148	1999
Volkswagen Werk Osnabrück	VW	2752	2009
Volkswagen Braunschweig	VW	7048	1938
Volkswagen Salzgitter	VW	7179	1969
Volkswagen Zwickau	VW	8419	1990
Volkswagen Chemnitz	VW	1789	1991

Table B.2
Local Firms - Industries

This table lists the WZ 2008 industry codes and industry names that we use to define local firms.

code	name
47	Retail trade, except of motor vehicles and motorcycles
55	Accommodation
56	Food and beverage service activities
62	Computer programming, consultancy and related activities
63	Information service activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
62	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
82	Office administrative, office support and other business