DIW BERLIN

1794

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2019

Occupational Characteristics and the Gender Pay Gap

Aline Zucco

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Occupational characteristics and the Gender Pay Gap*

Aline Zucco

DIW Berlin, Research Group Gender Studies,
Phone: +49 30 89789-383, E-mail: AZucco@diw.de

Abstract

Germany has a large persistent Gender Pay Gap of 21 %; although this gap is not constant across occupations. The question arises why some occupations have large Gender Pay Gaps while others have only small gaps. Using data from the Structural Earnings Study merged with occupational task information provided by the Federal Labor Office, this paper aims to uncover the relationship between occupational characteristics and the Gender Pay Gap. To do so, I apply a two-step approach, where the first step uses individual characteristics to estimate the adjusted occupation-specific Gender Pay Gaps. In the second step, these gaps are regressed on occupational characteristics. I find that wage differences between men and women are lower in occupations with linear earnings and in occupations with a large share of public firms. Moreover, we observe that an increasing share of persons with supervisory power is linked to larger wage differences between men and women, which indicates the presence of a glass ceiling. Finally, the Gender Pay Gap is higher in occupations with routine tasks. Moreover, the findings suggest that the more that employees can be substituted with other employees, the lower is the Gender Pay Gap. Hence, this study extends previous findings on occupation-specific Gender Pay Gaps by linking them to occupational characteristics on a more general level.

JEL Classification: J3, J31, J24, J16
Keywords: Gender pay gap, segregation, discrimination, wage differentials, occupations

*I thank Peter Haan, Katharina Wrohlich, Anna Hammerschmid, Julia Schmieder, Sophia Schmitz and the participants at the 2017 BeNA seminar and at the 2018 workshop for feminist economics for valuable comments and fruitful discussions. Further, I am very grateful to the team at the Research Data Center of the Federal Statistical Office.

1 Introduction

In 2017, the German Federal Statistical Office reported that the raw Gender Pay Gap (GPG), i.e. the relative wage differences between men and women, is 21 % in Germany. The report shows that occupational choice plays an important role in the GPG (Finke, Dumpert, and Beck, 2017). As shown in table 1, which presents the GPGs in the ten largest occupations in Germany, the GPG varies substantially across occupations: For instance, in Nursing as well as Education and Social Work, women and men have similar wages on average. In contrast, in occupations such as Machine-building and -operating, and Business Organization, the GPGs are very large. In these occupations men earn on average 25 % and 35 %, respectively, more than women. Hence, the question arises why the GPGs vary so much between occupations and how occupational characteristics are related to these differences.

Occupation	Gender Pay Gap
Education and Social Work	0.02
Nursing	0.02
Cleaning Services	0.07
Warehousing and Logistics	0.17
Office Clerks and Secretaries	0.18
Drivers of Vehicles in Road Traffic	0.18
Public Administration	0.20
Average	0.21
Sales Occupations (without Specialization)	0.24
Machine-building and -operating	0.25
Business Organization	0.35

Source: SES 2014, values are weighted

Table 1: Gender Pay Gap in the ten largest occupations in Germany

In this paper, I systematize occupational differences in order to reveal whether, and to which extent, the GPG is linked to occupational characteristics in Germany. Using the Structure of Earnings Study (SES) data from the Federal Statistical Office matched with task information provided by the Institute of Employment Research (IAB), I link individual and occupational characteristics with hourly wages. Applying a two-step-approach, I first estimate the GPGs within occupations. Second, I descriptively systematize the dif-

ferences between occupations by explaining the variance in the occupation-specific GPGs. For this systematization, I aggregate information on working conditions at the occupational level.

To highlight the relationship between occupational characteristics and the GPG, I show differences in the GPG between different occupations. In the raw data, we observe higher GPGs in occuppations with mainly leadership positions and in occupations with interactive and analytical non-routine tasks. Moreover, the data suggest a relationship between the linearity in earnings and the size of the GPG. An occupation is defined as linear if hourly wages are constant along the distribution of working hours. In contrast, persons who are employed in occupations with non-linear earnings face wage premia for longer working hours. Specifically, the data reveal that in particular occupations in the medical sector, that have low GPGs, tend to remunerate linearily. In contrast, occupations with non-linear earnings, which are more pronounced in the business sector, have higher GPGs. However, these findings may result from differences in observables.

Therefore, I estimate, in the first step, the adjusted GPG within occupations based on individual characteristics. After controlling for human capital and firm characteristics, the average GPG within occupations is 13 %. While in some occupations women earn more than men (e.g. Civil Engineering or Event Organization), female employees in Legal Services earn 33 % less, while and Actresses, female Dancers and Athletes earn even 53 % less than their male colleagues. Excluding part-time workers from the regression leads to slightly different results, which emphasizes the impact of part-time workers when estimating the GPG.

In the second step, I regress occupational characteristics on the GPG obtained in the first step. To measure, whether or not, the linearity in earnings is related to the GPG, I introduce the non-linearity index. This index gives the relative occupation-specific difference in the hourly wage between persons working more than 40 hours per week and those working less than 25 hours. I find that occupations with more linear earnings show more equal wages between men and women. Moreover, in contrast to the raw data, occupations with more routine tasks have larger GPGs on average. According to the literature, occupations with linear earnings (Goldin, 2014) and with non-routine tasks (Bhalotra and Fernández, 2018) have a higher level of substitution. These relations indicate the importance of substitution when it comes to the GPG: The more that employees can replace each other, the less pricey is the absence of a particular employee and the lower is the GPG.

Moreover, GPGs are higher in occupations where a high share of employees have supervisory power, which indicates a glass ceiling. In addition, I find that the share of public firms

reduces the inequality in wages between men and women as these firms are more likely to provide collective agreements. These agreements do not just decrease the leeway in discrimination among workers but might also have positive external effects on other firms. Hence, private companies may reward their employees according to the wage agreements of the public firms.

A considerable literature examines the various reasons for the GPG.[1] Besides the large strand that focuses on gender differences on behavior (Babcock and Leschever, 2003; Bertrand, 2011; Croson and Gneezy, 2009; Niederle and Vesterlund, 2007), many earlier studies focus on selection. That is, men often earn more than women because they select in better paying firms (Card, Cardoso, and Kline, 2016; Coudin, Maillard, and Tô, 2018; Goldin, Kerr, Olivetti, and Barth, 2017) or working in occupations with higher earnings (Blau and Kahn, 2017; Ludsteck, 2014; Murphy and Oesch, 2016). But even within occupations, there is still a substantial GPG (Goldin, 2014; Hinz and Gartner, 2005). However, less is known on the occupation-specific GPGs, and more precisely, why they vary substantially across occupations.

My interest on the role of occupational characteristics to explain differences in the GPGs between occupations is based on a study by Goldin (2014). In this paper, she shows that the GPGs in the American labor market vary substantially between occupations and that is linked to the degree to which hourly wages increase with the number of working hours. Hence, in occupations where the wage level is independent of working hours, the GPG is lower than in those occupations where earnings increase disproportionately with the number of hours worked.

There is still uncertainty, however, to what extent these results can be transferred to other labor markets. In this paper, I focus on the German labor market because it is characterized by a high share of part-time work. However, part-time work is a quite female phenomenon: in 2017, 48 % of women and 11% of men worked in part-time (Statistisches Bundesamt, 2018a). Moreover, the role of occupations in Germany is very important as they determine, to a high degree, the professional pathway. Since the German education and vocational training system is highly standardized, apprenticeship training serves as a strong signal for a specific knowledge in one particular occupation. As a result, the number of occupational shifts decreases, while making occupational changes rather complicated (Allmendinger, 1989).

Moreover, I extend the analyses of Goldin (2014) first by introducing a non-linearity index that allows to show a more general link between the linearity in earnings and the GPG.

[1] A large review of the current state of research is given by Blau and Kahn (2017).

Second, I incorporate additional characteristics, such as the distribution of hierarchy levels and the tasks on the occupational level to describe why the GPGs vary between occupations.

The paper is structured as follows. The next section presents some bivariate correlations between the raw GPG and some occupational characteristics. The third section is concerned with definitions and the data. Section four describes the estimation strategy used to analyze the relationship of occupational characteristics and the GPG, which consists of two steps. In the first step, I estimate the occupation-specific GPGs. The second steps aims to systematize the differences in the gaps across occupations by regressing the GPG on occupationals characteristics. Section five provides the results of the estimation, while section six discusses the findings. The last section concludes.

2 Why the GPG may be linked to occupational characteristics

Before discussing the empirical strategy and results, this section presents bivariate correlations between wages and several occupational characteristics, which highlights the role of occupations when analyzing the GPG.

2.1 Linearity of earnings

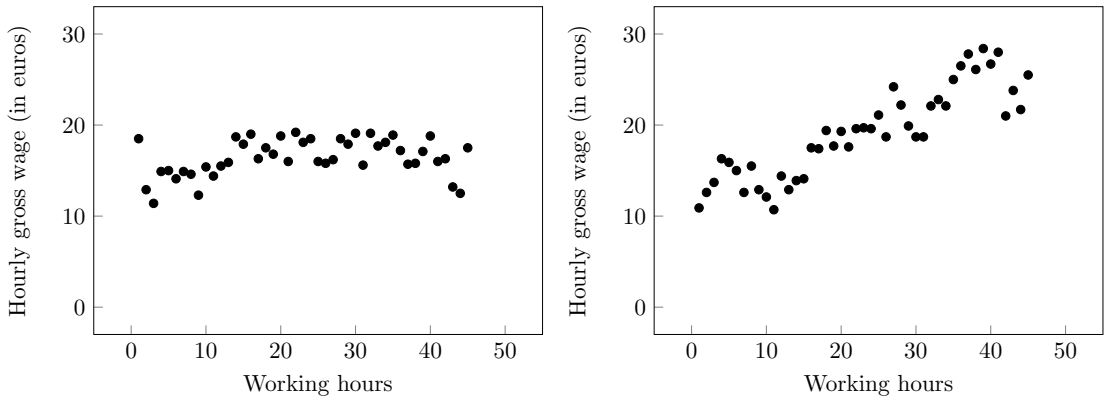
For the American Labor Market, Goldin (2014) shows that differences in the GPGs between occupations are related to the degree to which hourly wages depend on the number of working hours. In this context, she differs between occupations with "linear" and "non-linear" earnings. In occupations with linear remuneration, the hourly wages are independent from the number of hours worked and thus, earnings increase linearly with the working hours. In contrast, in occupations with non-linear or convex earnings, wages rise with the number of working hours. Therefore, the earnings increase disproportionately with the hours worked. Goldin (2014) argues that in occupations with non-linear remuneration presence is of high value and therefore, flexible working hours is costly to the firm as employees are not available at a specific time. Conversely, workers in occupations with linear earnings can easily be substituted by each other such that flexible working hours do not lead to higher costs for the employers.

She observes that occupations with linear earnings (e.g. pharmacy) have lower GPGs than those with non-linear earnings (e.g. MBA, JD). As part-time workers are predominantly female, the (non-)linearity of earnings can partly explain why the GPG varies over occupations in the U.S.

The correlation between the linearity of earnings and the size of the GPG may also hold true in the German labor market, as shown in figure 1. Out of the ten largest occupations in Germany, Business Organization is the one with the highest GPG, while Nursing has the smallest (see table 1). Comparing the size of the gross hourly net wages in both occupations by the number of weekly working hours, shows notable differences the correlation of working hours and wages. Employees in Nursing have, on average, the same wage independent of the numbers of working hours. The average gross wages of persons working 20, 30, or 40 hours is 19 euro per hour. Hence, the remuneration in these occupations is linear.

In contrast, the average gross wages in Business Organization increases with the number of working hours per week. That is, employees with 15 hours earn 14 euro, those with 30 hours earn 19 euro and those with 40 hours earn 27 euro per hour. As salaries rise disproportionately with the hours worked, Business Organization is defined as an occupation with non-linear remuneration.

These examples show that the wage level is in some occupations more dependent on the number of working hours than in others. Moreover, occupations with non-linear earnings tend to have a higher GPG than occupations with linear earnings. This correlation indicates that the degree of linearity in earnings may be related to the differences in the GPGs between occupations.



(a) Nursing

(b) Business Organization

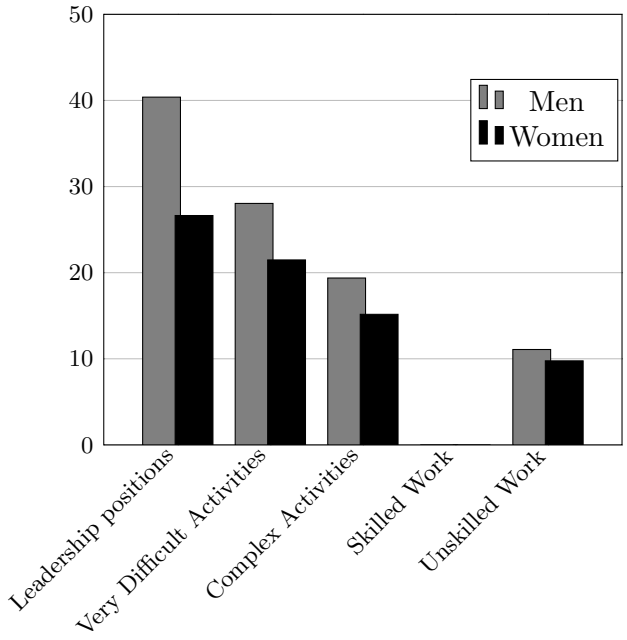
Source: SES 2014; Based on all employees between 25 and 55 years.
Values are weighted.

Figure 1: Hourly gross wages by working hours for employed in Business Organization and in Nursing

2.2 Hierarchy and tasks

Moreover, a considerable literature shows that the GPG is substantially large at the top of the wage distribution (Arulampalam, Booth, and Bryan, 2007; Blau and Kahn, 2017; Busch and Holst, 2009; Collischon, n.d.; Gallego Granados and Wrohlich, 2018). This finding may be the result of a "glass ceiling," i.e. it is difficult for women to enter top positions.

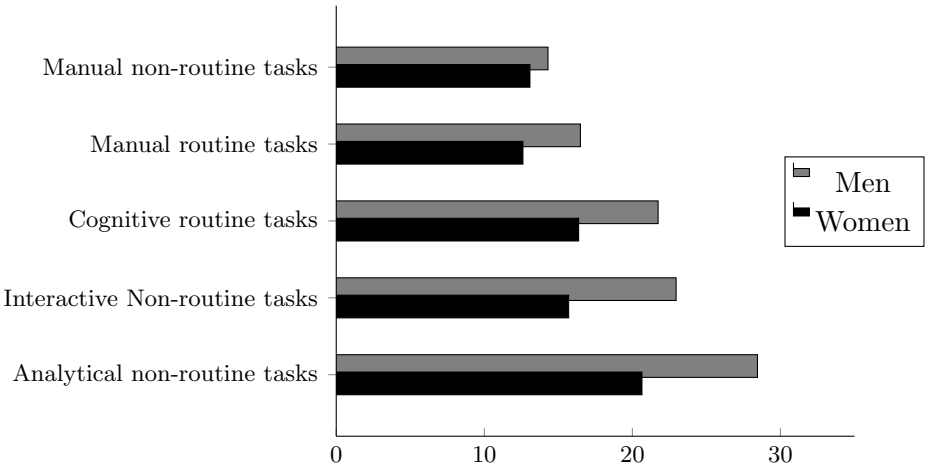
Hence, the variation in the GPG between occupations may be related to the fact that occupations are differently affected by the glass ceiling. To test this relationship, I make use of the "performance group" that describes the hierarchical rank of each employee.



Source: SES 2014; Based on all employees between 25 and 55 years.
Values are weighted.

Figure 2: Wage level within hierarchical groups by sex

Figure 2 classifies occupations by the mode of the hierarchical groups, giving the average male and female wage level for occupations that are mostly characterized by leadership positions, very difficult activities, complex activities or unskilled work. None of the occupations is mainly defined by skilled work. The graph shows that wage differences between men and women increase with the hierarchal level. The largest gaps occur in occupations with mainly leadership positions. These descriptive results suggest the existence of a glass ceiling. As some occupations may be more affected by a glass ceiling than others, the share of leadership positions within an occupation may be correlated with the GPG.



Source: SES 2014; Based on all employees between 25 and 55 years. Values are weighted.

Figure 3: Wage level within main tasks by sex

In addition, the literature emphasizes the role of tasks to explain the GPG. Black and Spitz-Oener (2010) show that the decrease of the GPG over time is partly related to changes in the work content because of workplace computerization. This is why, the share of non-routine interactive and analytical tasks has increased more for women than for men. In contrast, women's share of routine tasks has decreased stronger than men's. Moreover, computerization decreases the relative price of routine tasks. Thus, task-based technological change favors women more than men, and is therefore partly explaining why the GPG has decreased over time.

However, less is known about GPGs within tasks. Therefore, occupations are grouped on the task that is mainly performed, with figure 3 showing male and female wages within each task. The graph indicates that men's wages differ substantially across tasks: While men, earn on average, 14 euro per hour in occupations with manual non-routine tasks, the male wage level in occupations with analytical non-routine tasks is around 28 euro per hour. In contrast, female wages are more constant across tasks, varying between 12 and 16 euro. Only in occupations with analytical non-routine tasks is the average hourly wage above 20 euro.

Hence, figure 3 illustrates that the GPG varies across tasks. Moreover, it indicates that these gaps between men and women are mostly related to remarkable heterogeneities of male wages between different tasks.

To sum up, comparing wages between occupations with different occupational characteristics suggests that they may be correlated with the GPG. However, these findings are based on simple group comparisons and may also result from differences in observables. Therefore, the next sections provide more information on the data set and the empirical strategy to test whether, and if so, which occupational characteristics are linked with the adjusted GPG.

3 Data and Descriptives

3.1 Data Source

The estimation is based on the SES, which is a linked employer-employee data set provided by the Federal Statistical Office. The data set offers detailed information on work characteristics, including earnings and hours worked. The data come from the employers or, in Education and Public Administration and Defense or Social Insurance sectors, from the personnel statistics of the public service (Statistisches Bundesamt, 2018c). In contrast to survey data, the administrative provision of wage information substantially decreases the likelihood of measurement errors. Due to the duty of disclosure, nonresponse, which is often a concern in survey data, does not bias the results (Kapteyn and Ypma, 2007).

The SES are cross-sectional data that are collected every four years since 2006. The data offers information about the employee (e.g. gender and occupation), the employment (e.g. wage and working hours) and the employer (e.g. firm size and public vs. private ownership).

In this paper, I use the 2014 wave, which also provides information on performance group, shift work, leadership positions, and overtime hours. The gross sample size of employers exceeds 60,000, while that of employees exceeds 1 million observations (Statistisches Bundesamt, 2018c). This large sample size is a major benefit of the data as it allows detailed analyses within occupations.

Another important advantage of the data is that it provides information on working hours. Existing studies (e.g. Gartner and Hinz, 2009; Hirsch, 2013) that estimate the size of the GPG in Germany use administrative data provided by the IAB. As IAB data offer daily wage information but not work hours, the analyses are usually only based on full-time employed persons (e.g. Gartner and Hinz, 2009) to make wages more comparable. Re-

stricting the sample to full-time employees not only excludes nearly half of the females[2] but also concentrates on a very selective group of women.

3.2 Definitions

i. Hourly wages

The estimation of the GPG is based on hourly wages, which relies on the number of agreed working hours per week plus the number of paid over-time hours. Further, this number is multiplied times 4.3 to determine the number of agreed monthly hours. Finally, the monthly gross earnings, which includes pay for overtime and shift work, is divided by the number of working hours per month.

ii. Occupations

Occupations form the key element of this study and group similar jobs with similar formal training. They are defined based on the three-digit-level[3] and differentiate between 144 occupations, which are given in the appendix. As an example, this definition allows for distinguishing between human and veterinary medicine, but not between surgeons and pediatricians.

iii. Tasks

The SES are merged with aggregated data on tasks for 2013[4] provided by the IAB. This data contain the composition of tasks and the main task within each occupation (see Dengler, Matthes, and Paulus (2014) for more detail). The tasks are grouped in the following way: Analytically non-routine tasks, interactive non-routine tasks, cognitive routine tasks, manual routine tasks and manual non-routine tasks. An overview of the specific activities within each tasks is presented in table (6) in the appendix.

As an example, in the occupation of Education and Social Work, the share of analytically non-routine tasks is 33 %, of interactive non-routine tasks is 51 %, and of manual non-routine tasks is 11 %, while the share of routine tasks in this occupation is rather small. Only 5 % are cognitive-routine tasks and there exist no routine manual tasks. In contrast, industrial occupations such as glass- or ceramic-making have mainly routine manual tasks. The task-data does not offer information on soldiers. For this reason, I exclude four occupations, which describes different ranks of the German army, from the analysis.

[2] As mentioned before, the 48 % of the employed women work in part-time (Statistisches Bundesamt, 2018a).

[3] The assignment to the different occupations is based on the classification of occupations 2010 (KldB 2010, Paulus and Matthes (2013)).

[4] I assume that tasks within occupations are stable over time.

iv. Non-linearity index

To measure the linearity of earnings within occupations, I introduce a non-linearity index. It gives the relative wage gap within occupations between persons working more than 40 hours per week and those working less than 25 hours. The less linear the remuneration within an occupation, the larger is the hourly wage gap between persons working more than 40 and less than 25 hours, and, thus, the higher is the index.

Table 2 shows the non-linearity index for the five occupations with the most (non-)linear remuneration given that the occupation has more than 1,000 observations. In these "most linear" occupations, such as sales occupations for drugstore products or journalism, part-time workers earn more than those with more than 40 hours/week. In contrast, managers who work less than 25 hours/ week earn 47 % less than their colleagues with long working hours.

In comparison to the American labor market, Germany shows similar trends with regard to occupations with (non-)linear earnings: Occupations in the medical sector, such as selling drugstore products or pharmaceuticals or nursing, tend to remunerate more linear, while occupations in the business sector such as business organization or managing are occupations with highly non-linear earnings. This finding emphasizes that despite the international differences between these labor markets, the conditions within occupations seem to be similar.

Occupation	Linearity (in %)
Sales occupations (Drugstore products, pharmaceuticals)	−7.8
Occupations in editorial work and journalism	−3.8
Occupations in civil engineering	−2.4
Doctors' receptionists and assistants	−1.6
Occupations in nursing and obstetrics	−1.3
⋮	⋮
Occupations in security and safety	35.7
Occupations in business organization and strategy	35.8
Occupations in building services and waste disposal	36.8
Technical occupations in paper-making and packaging	38.3
Managing directors and executive board members	47.0

Source: SES 2014;
Note: Linearity index in occupations with more than 1,000 observations;

Table 2: Occupations with the most (non-)linear earnings

3.3 Descriptive Overview

I restrict the sample to persons between 25 and 55 years of age and exclude trainees and those working less than 9 hours per week. The final sample contains 434,821 employees, of which 194,608 are women. Table 3 provides the summary statistics for the entire sample and separately for women and men. With respect to their individual characteristics men and women differ in their wages, tenure, and the probability of holding a leadership position. As expected, men have significantly higher hourly wages than women. While men earn on average 20.9 € per hour, the average hourly wage of women is more than 5 € smaller. Moreover, men tend to work longer in the same establishment than women and are more likely to hold a leadership position. Additionally, the average age is 41 years and 88 % of the sample has a permanent contract.

The majority of the sample (63 %) does not have any A-Levels; instead they have vocational training, with a distinct minority of the sample having completed tertiary education (University: 14 %, Polytechnical school: 2 %). The share of persons without vocational training is slightly smaller than the official numbers in the microcensus provided by the federal statistical office. As the sample is restricted to the working population, it is better educated than the average population.

One quarter of the (female and male) employees work in East Germany[5]. Furthermore, 72 % of the establishments are located in urban regions. On average, the establishment have more than 700 and companies more than 3,700 employees but men tend to work in larger establishments and companies.

The third part of the table shows occupational characteristics. The table demonstrates that women work in occupations with fewer overtime hours and smaller shift bonuses. In contrast, men and women work in occupations with similar level of linearity of earnings.

Regarding the distribution of hierarchy levels, typically employees work in occupations, where the majority has difficult activities and the minority does unskilled work. The distribution of these groups does not differ between men and women.
Moreover, the table shows how tasks are distributed across the sample. Employees work mostly in occupations with cognitive routine (30 %) or analytical non-routine tasks (23 %). In addition, men are more likely to work in occupations with manual tasks and women in occupations with more interactive non-routine tasks.

[5]East Germany comprises the new federal states (former GDR), while Berlin and West Germany make up the old federal states (former FRG)

	Total		**Women**		**Men**	
	Mean	SD	Mean	SD	Mean	SD
			Individual Characteristics			
Hourly wage	18.61	11.93	15.81	8.51	20.89	13.69
Female	0.48	0.50	1	0	0	0
Age	41.21	8.96	41.26	9.04	41.17	8.88
Tenure	9.14	8.70	8.51	8.31	9.64	8.97
Education						
No A-level, No VT	0.07	0.25	0.07	0.26	0.06	0.24
No A-Level, VT	0.63	0.48	0.61	0.49	0.64	0.48
A-Level, No VT	0.03	0.17	0.03	0.17	0.03	0.17
A-Level, VT	0.12	0.32	0.14	0.35	0.10	0.30
Polytechnical degree	0.02	0.14	0.02	0.14	0.02	0.14
University	0.14	0.35	0.12	0.33	0.15	0.36
Permanent Contract (vs. Temporary Contract)	0.88	0.33	0.86	0.35	0.89	0.31
Leadership Position	0.06	0.25	0.04	0.19	0.09	0.28
			Firm Characteristics			
East vs. West Germany	0.23	0.42	0.23	0.42	0.23	0.42
Size of the Establishment	752.94	2967.02	541.19	2036.83	924.49	3536.72
Size of the Company	3757.96	17,321.04	2814.93	13,754.35	4521.96	19,710.64
Urban Region (vs. Rural Region)	0.72	0.45	0.72	0.45	0.72	0.45
			Occupational Characteristics			
Overtime hours	1.55	1.40	0.97	0.90	2.02	1.55
Shift bonus	42.27	52.09	28.70	42.08	53.27	56.63
Non-linearity Index	0.22	0.11	0.20	0.12	0.24	0.11
Hierarchical groups						
Supervisory Power	0.09	0.14	0.09	0.13	0.10	0.14
Very Difficult activity	0.20	0.13	0.19	0.13	0.20	0.14
Difficult activity	0.53	0.17	0.54	0.19	0.53	0.16
Skilled work	0.12	0.10	0.12	0.10	0.12	0.10
Unskilled work	0.06	0.09	0.06	0.12	0.05	0.07
Tasks						
Analytical Non-Routine Tasks	0.23	0.42	0.20	0.40	0.26	0.44
Interactive Non-Routine Tasks	0.12	0.33	0.20	0.40	0.06	0.24
Cognitive Routine Tasks	0.30	0.46	0.30	0.46	0.29	0.46
Manual Routine Tasks	0.14	0.35	0.07	0.26	0.20	0.40
Manual Non-Routine Tasks	0.21	0.41	0.24	0.42	0.19	0.39
Share of Women	0.45	0.29	0.63	0.23	0.30	0.25
Public Firms	0.15	0.18	0.18	0.20	0.12	0.15
N	434,821		194,608		240,213	

Source: SES 2014;

Table 3: Summary Statistics

As expected, the summary statistics reveal that women tend to work in female occupations and men in male occupations. Further, women are more likely to work in public firms.

4 Empirical Strategy

Section 2 presents binary relations between occupational characteristics and the raw GPG. However, parts of these wage differences may come from dissimilarities regarding education levels, tenure or firm size. To control for these differences, I estimate the adjusted GPG within occupations as a first step. In the second step, I use occupational characteristics to systematize the variance in the adjusted GPGs between occupations.

Thus, as a first step, hourly wages are regressed on age, tenure, education, location of the establishment (East vs. West Germany, urban vs. rural area) size of the establishment and the company, having a leadership position, and having a permanent contract. These variables are summarized in vector X_i. In addition, the model contains gender (δ_i), occupational fixed effects (γ_j), and their interaction (α_j). This interaction term gives the conditional wage (y_{ij}) differences between men and women within each occupation j, i.e. the adjusted GPG, and, therefore, is the coefficient of interest. The indicators illustrate that each individual i is working in an occupation j.

$$log(y_{ij}) = \delta_i \; Female_i + \sum_{j=1}^{J} \gamma_j \; Occ_j + \sum_{j=1}^{J} \alpha_j \; Occ_j * Female_i + \mu X_i + \epsilon_{ij} \qquad (1)$$

However, it is likely that some unobserved preferences are correlated with the explanatory variables, especially with the occupation fixed effects. Hence, persons working as managers may have stronger preferences for professional success, while employees in Education and Social Work may seek a better reconciliation of family and working life. That is, wage differences between those two occupation may also come from selection processes.

In addition, the α_j coefficient might be biased due to different selection processes within occupations. It is possible that in some occupations, such as medicine, men and women select into different occupational sub-groups with diverging wage levels. But women are more likely to select into the relatively low remunerated occupational sub-group of pediatricians as into the sub-group of surgeons, which has a relatively high wage level (Gesundheitsberichterstattung des Bundes, 2019; Statistisches Bundesamt, 2018b). Therefore, the size of α_j might not necessarily come from discrimination, but might also be the result of (unobserved) selection.

Thus, the estimates of (1) would be unbiased if the residuals were orthogonal. But due to the selection within occupations and occupational groups, the condition formulated in

(2) might be violated. Therefore, the estimated coefficients do not necessarily reflect the causal link between the independent variables and the wage.

$$\mathbb{E}(\epsilon_j | Female_i, Occ_j, Occ_j * Female_i, X_i) = 0 \tag{2}$$

The second step aims to systematize the variance in the adjusted GPG, α_j, between occupations. Therefore, the occupation-specific wage differences, estimated in (1), are regressed on different occupational characteristics (δ) such as shift bonus, non-linearity of earnings, and tasks.

$$\hat{\alpha}_j = \beta_0 + \delta C_j + \nu_j \tag{3}$$

In equation (3), the selection into occupations also has a crucial impact on the interpretation of the results because it affects both sides of the equation. As mentioned above, it is probable that estimated α_j differs from the true α_j^*. Therefore, the relationship between the true and the estimated coefficient can be formulated as following, where a signifies the measurement error:

$$\alpha_j = \alpha_j^* + a \quad \text{with} \quad a \neq 0 \tag{4}$$

Thus, we can rearrange equation (3) as follows:

$$\hat{\alpha}_j^* = \beta_0 + \delta C_j + a + \nu_j \tag{5}$$

The results of the estimation are only biased if the measurement error a is correlated with the covariates. However, this is very likely, if we think, for example, about the share of women within an occupation. Hence, women working in a female-dominated occupation, such as sales or nursing, may prefer to work in these occupations as they offer a better reconciliation of family and working life due to flexible working hours. Men, in contrast, may work in these occupations for other reasons. It is also possible that women working in male-dominated occupations, like technical occupations in the automotive building industry, may be a very selective group. That is, these women might be more ambitious than the average population and, therefore, also more labor market attached than men in these occupations.

In addition, the selection into occupations also may affect the right-hand side of equation (5). It is possible that persons may select into an occupation because it has shift bonuses or, in contrast, other persons may not work in this occupation because it includes shift work.

Because people do not select randomly into occupations, the assumption of orthogonal residuals does not hold (equation (6) and (7)).

$$\mathbb{E}(\nu_j | C_j) = 0 \tag{6}$$

$$\mathbb{E}(a|C_j) = 0 \tag{7}$$

As a result, the estimates of the second step (3) are likely to be biased. This is why the coefficients cannot be interpreted in a causal way. However, even if the results come from selection, the estimated coefficients tell us more on wage differences between women and men: The results indicate whether, and if so, which occupational characteristics are linked to the GPG.

5 Results

This section presents the relationship between occupational characteristics and the GPG. First, I present the estimation results of occupation-specific GPGs. In the second step, I descriptively systematize the differences in the adjusted GPGs between occupations. Therefore, I analyze to what extent occupational characteristics can explain the variance in the GPGs.

5.1 First step: The GPG within occupations

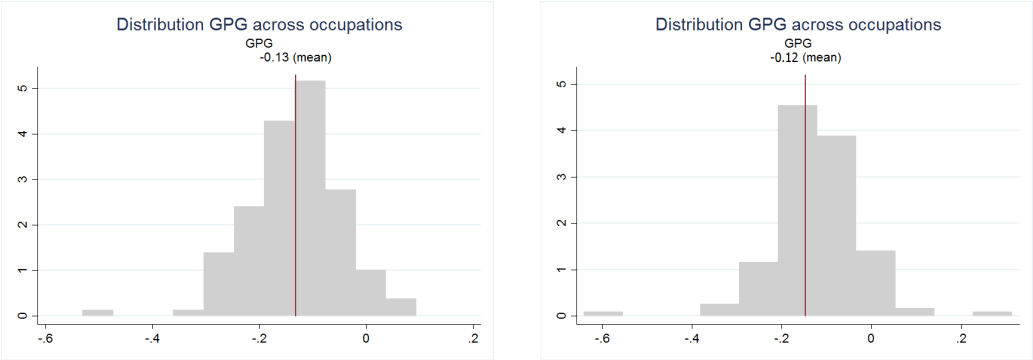
Figure (4) graphs the distribution of the coefficient of interest α_j, estimated in equation (1), that gives the adjusted GPG in each occupation for two specification.[6] The first specification shows how the GPGs are distributed if equation (1) is estimated for the entire sample (a). The second specification is based on full-time employees only (b).

The results show that there are substantial differences in the adjusted GPGs between occupations. Moreover, the graphs emphasize that the distribution of the gaps changes slightly if part-time worker are excluded. The GPGs in the full-time sample tend to be slightly smaller. This may be the result of a positive selection of women working in full-time (Gallego Granados, 2019).

In addition, the sample selection could affect the estimation results of previous studies (e.g. Ludsteck, 2014), which estimated the occupation specific GPGs based on full-time employees. The results would be biased if the relationship between the occupations, thus the ranking, is affected. Therefore, table 4 presents the occupations with the highest and lowest GPG in both specifications.

In the main specification, which is based on the entire sample, Event Organization and Alternative Medicine have the lowest GPG, which is even positive. That is, in these occupation men earn on average 6 % and 4 % less than women. There are other occupations

[6] The estimation results of equation (1) for the entire and the full-time sample is given in the appendix in table (8)



(a) Entire Sample (b) Full-time Sample

Source: SES 2014; GPG equals α_j

Figure 4: Distribution of the Gender Pay Gap

like Civil Engineering, Education and Social Work, as well as Church Community Work, where women earn more than men. Note that the raw GPG for persons employed in Education and Social Work (see table 1) turns from a slightly negative to a positive value after controlling for observables. Thus, in this occupation the (raw) GPG may be rooted in from different human capital endowments or the fact that men are more likely to hold leadership positions.

However, in other occupations, like Managing Directors and Legal Services, women earn 26 % and 33 % less than men, which is similar to the American labor market (Goldin, 2014). The highest GPG occurs for Actors, Dancers and Athletes, where women earn 51 % less than men. However, these are a rather small occupations, where a few outliers can have an outsized impact on the results.

In the second specification, which excludes part-time worker, the ranking changes slightly. Hence, the selection into full-time might affect occupations differently. This results highlights the importance of adding part-time workers when comparing GPGs across occupations.

5.2 Second step: The relationship between occupational characteristics and the GPG

The second step aims to explain the variance of the GPG across occupations that was obtained in the first step. Therefore, I add occupational characteristics stepwise to the model.

	Entire Sample		Only Full-time	
1	0.06	Event Organization	0.08	Vini- and Viticulture
2	0.04	Alternative Medicine	0.05	Civil Engineering
3	0.03	Civil Engineering	0.05	Event Organization
4	0.01	Church Community Work	0.04	Occupational Health and Safety
5	0.01	Education and Social Work	0.03	Education and Social Work
6	0.01	Service in Passenger Traffic	0.03	Church Community Work
7	−0.01	Presenters and Entertainers	0.01	Underground and Surface Mining
8	−0.02	Occupational Health and Safety	0.01	Special Interest Organizations
9	−0.02	Sales Occupation: Books and Art	0.00	Drivers of Construction Vehicles
10	−0.03	Nursing	0.00	Floristry
		⋮		⋮
131	−0.25	Metal-Making	−0.24	Sales Occupations: Durables
132	−0.26	Managing directors	−0.24	Artisans designing ceramics
133	−0.26	Fishing	−0.24	Musical Instrument Making
134	−0.26	Product and Industrial Design	−0.24	Paper-making and -processing
135	−0.27	Photography	−0.25	Photography
136	−0.27	Sales occupations: Durables	−0.26	Printing Technology
137	−0.29	Printing Technology	−0.29	Sales Occupations: Foodstuffs
138	−0.30	Paper-making and -processing	−0.29	Product and Industrial Design
139	−0.33	Legal Services	−0.34	Legal Services
140	−0.51	Actors, Dancers and Athletes	−0.64	Actors, Dancers and Athletes

Source: SES 2014; The GPG equals the exponential value of α_j minus 1.

Table 4: Occupations with the highest and the lowest Gender Pay Gap

The first row of table (5) includes the average number of overtime hours within occupations. The insignificance of the coefficient indicates that the differences in the GPGs between occupations are not correlated with the number of overtime hours. The summary statistics (see table 3) reveal that men work in occupations with higher over-time hours. This finding suggests that women are less likely to select into occupations with many over-time hours, but once they work in these occupations they do not earn less then men.

The second row shows that the shift bonus amounts correlate slightly negatively with the GPG. That is, the GPG is higher in occupations where shift work has a larger impact on wages. Shift bonus are paid for night work or for work on Sundays and holidays. Assuming that women are more concerned about reconciliation of family and working life (e.g. Blau and Kahn, 2017), they will be less likely than men to work on weekends or at night.

In the third row, the non-linearity index is included in the model. The negative and significant coefficient emphasizes the relationship between linearity in earnings and the

GPG, which was indicated in section two. Thus, the larger the relative wage difference between persons, who work less than 25 hours per week, and those working more than 40 hours, the higher is the occupation-specific GPG. As a reminder, women are more likely to work part-time and, therefore, are more affected by part-time penalties. This result confirms not only Goldin's (2014) findings but extends them to a more general level: The relationship between the linearity in earnings and the GPG, which is observed in some occupations such as Nursing or Business Organization, is not random but rather systematical.

However, this relationship might not (only) be the result of discrimination but can also come from selection into full- or part-time. In a recent study, Gallego Granados (2019) highlights that positive selection into full-time work has a substantial impact on the part-time wage gap. Hence, lower wages in part-time do not necessarily mean a part-time penalty but can also result from lower productivity.

In specification (IV), the distribution of hierarchal groups within occupations is added to the model. The results indicate that the GPG is greater in occupations with a large share of employees holding leadership positions, which is also observed in the raw data (section 2). As the link still holds after controlling for observables, this finding hints at the presence of a glass ceiling and, therefore, is in line with previous studies (e.g. Arulampalam et al., 2007; Collischon, n.d.). In addition, this result broadly supports the work of Busch and Holst (2009, 2011), which shows higher GPGs in managerial positions. Moreover, the insignificance of the coefficients of the remaining hierarchical groups reveals that the link between the GPG and the distribution of hierarchical groups is not linear but rather represents a penalty for women in leadership positions.

In addition to the glass ceiling, this relationship may also have other causes, such as differing negotiation skills. Previous research shows that women negotiate their wages less successfully than men (e.g. Niederle and Vesterlund, 2007) and negotiating occurs mainly in jobs with supervisory power.

Moreover, after including hierarchical groups, the coefficient of shift bonus becomes insignificant. This result indicates that the correlation between the shift bonus and the GPG in models (II) and (III) may be driven by occupations with large shift premia and high shares of employees in leadership positions.

Occupational tasks enter the model in the fifth row of table (5). The highly significant coefficients and the large increase of the R^2 emphasize the importance of tasks in explaining

differences in the GPG between occupations. The results reveal that the GPG is larger in occupations that have mainly cognitive or manual routine tasks.

The findings is consistent with that of Bhalotra and Fernández (2018), who observe a similar relationship for the Mexican labor market. The authors argue that the level of substitution between men and women, which varies across tasks, can explain the differences in the GPG. Compared to those with manual and routine tasks, occupations with analytical non-routine tasks have a high level of substitutability between the sexes, which thus decreases the GPG. If we assume that this link holds more generally, we can say that non-routine tasks have a higher level of substitution than routine tasks. Thus, this relationship underlies the argument of Goldin (2014) saying that the more that workers can replace each other, the lower is the GPG.

In contrast, higher GPGs in routine tasks may also result from selection. Adda, Dustmann, and Stevens (2017) assume that women with a higher preference for fertility select into occupations with manual routine tasks. That is, the GPG in these occupations may come from differences in labor market attachment between men and women.

The higher GPGs in routine tasks, however, is surprising if we compare it with the raw GPGs within tasks (figure 3). Without controlling for observables, we observe the largest (absolute) wage gaps in interactive and analytical non-routine tasks. These differences between the adjusted and the raw GPG indicates that women are either working in generally less remunerated occupations within these tasks or that men have higher endowments of human capital.

In rows (VI) and (VII), the linear and quadratic terms of the share of women is added to the model. The insignificance of both coefficients emphasizes that women earn less than men irrespective of whether they work in a female- or a male-dominated occupation. Like previous studies, we find that a large part of the GPG is linked to the fact that male-dominated occupations are on average higher rewarded than female-dominated occupations (e.g. Levanon, England, and Allison, 2009 in the U.S., Hausmann, Kleinert, and Leuze (2015) in Germany). However, once they are working in a male-or female-dominated occupation, the size of the occupation-specific GPG does not depend on the share of women.

The last specification includes the share of public owners within occupations to explain the GPG. The coefficient is highly significant, which indicates that public ownership is negatively correlated with the GPG. That is, the size of the GPG depends on whether an occupation is mainly performed in public or private firms. One possible reason for this

relationship may be that employees in the public sector are mainly paid in accordance with collective agreements. These contracts ensure that persons with the same work experience and educational degree earn the same, which prevents discrimination. In contrast, in the private sector wages are often negotiated. As women are typically less successful at wage negotiations than men (e.g. Croson and Gneezy, 2009), this might explain why the GPG is higher in the private sector.

In addition to collective agreements, the positive relationship between the share of public firms and the GPG may result from selection. Hence, it may be that more labor market orientated men prefer to select private firms.

Moreover, after including the firm ownership in the model, the coefficient of manual non-routine tasks becomes significant. That is, occupations with mainly manual non-routine tasks have, on average, higher GPGs. This finding is consistent with that of Bhalotra and Fernández (2018) who find lower level of substitutability between men and women in occupations with manual tasks. The result also highlights that this relationship is more likely to appear in occupations that are mainly done in private firms.

Thus, some characteristics like the linearity in earnings, the distribution of hierarchical groups, tasks, and the share of public owners are related to the size of the GPG. However, in total, occupational characteristics cannot even explain a third of the variance in the GPG. Hence, the major part of the GPG is not linked to the observed occupational characteristics; rather there might exist more characteristics that cannot be observed in the data. As an example, it is possible that the link between within-occupational segregation and the within-occupational GPG is stronger in some occupations than in others. In human medicine, for example, women are more likely to work as pediatricians, while men tend to work as surgeons; the latter having a higher average wage level (Gesundheitsberichterstattung des Bundes, 2019; Statistisches Bundesamt, 2018b). In the occupation of teachers, in contrast, men are more likely to be math or physics teacher and women are more likely to teach languages (Weeber and Hobler, 2015). This, however, has no effect on their wages as, due to collective agreements, these are independent of the subject taught.

GPG	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Overtime hours	−0.0072	−0.0025	−0.0015	−0.0015	0.0059	0.0064	0.0053	0.0087
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Shift bonus		−0.0003*	−0.0003*	−0.0002	−0.0002	−0.0002	−0.0002	−0.0002
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-linearity Index			−0.1413*	−0.1392**	−0.1377**	−0.1359**	−0.1288**	−0.1304**
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hierarchy level (Ref: Difficult activity)								
Supervisory Power				−0.1027*	−0.1461**	−0.1458**	−0.1328**	−0.1821**
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Very difficult activity				0.0392	0.0034	0.0045	0.0202	−0.0330
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Skilled work				−0.0920	−0.0370	−0.0371	−0.0156	−0.0076
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unskilled work				−0.0140	−0.0270	−0.0299	−0.0272	−0.0744
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tasks (Ref: Analytical non-routine tasks)								
Interactive non-routine tasks					0.0294	0.0277	0.0269	0.0293
					(0.000)	(0.000)	(0.000)	(0.000)
Cognitive routine tasks					−0.0623**	−0.0625**	−0.0666**	−0.0638**
					(0.000)	(0.000)	(0.000)	(0.000)
Manual routine tasks					−0.0964**	−0.0964**	−0.0960**	−0.0916**
					(0.000)	(0.000)	(0.000)	(0.000)
Manual non-routine tasks*					−0.0422	−0.0427	−0.0465	−0.0528
					(0.000)	(0.000)	(0.000)	(0.000)
Share of women						0.0068	−0.1098	−0.1255
						(0.000)	(0.000)	(0.000)
Share of women^2							0.1305	0.1429
							(0.000)	(0.000)
Public firms								0.1371***
								(0.000)
Constant	−0.1352	−0.1304	−0.1045	−0.0929	−0.0625	−0.0667	−0.0566	−0.0627
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Occupations	140	140	140	140	140	140	140	140
R^2	0.0044	0.0174	0.0712	0.0686	0.1472	0.1407	0.1400	0.2136
R^2_{adj}	0.0115	0.0315	0.0912	0.1155	0.2147	0.2149	0.2204	0.2928

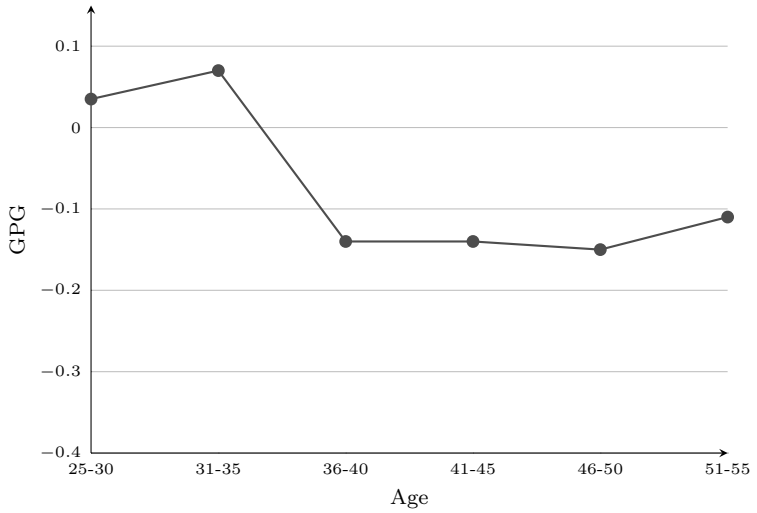
Source: SES 2014;

Note: GPG equals α_j obtained in the first step; Standard errors presented in parentheses; Significance levels: * p<.1; ** p<.05; *** p<.001

Table 5: Results: Second step

6 Discussion: What the data cannot explain

The previous section shows that some occupational characteristics are linked with the GPG but the question arises as to why do we still observe substantial differences in the gaps between occupations. To answer this question, it might be useful to take a step back and to look at wage differences in greater detail. By graphing the GPG across age groups, figure 5,[7] highlights that the wage differences between men and women varies across age. While women aged 25 to 35 earn even more than their male colleagues, the GPG decreases substantially in the age category 36 to 40 to -14 % and remains around this level until the age of 55, which is consistent with previous research (e.g. Goldin, 2014).



Source: SES 2014; Based on all employees between 25 and 55 years.

Figure 5: Adjusted wage gaps within age groups

To see whether or not this trend differs between occupations, figure 6 plots the adjusted GPG in three occupations with a high GPG (Machine-Building and -Operating, Sales Occupation (Foodstuffs), and Business Organization) and one with no GPG (Education and Social Work). The graph emphasizes that the development of the gap differs substantially across occupations. To test whether these dissimilarities may also come from group selection or different employment biographies, figure 7 graphs the share of women within occupations for the full-time sample and the entire sample.

[7]The model was extended by age and age fixed effects and age-female interactions. For occupation-specific coefficients, the model includes age-occupation fixed effects and respective interactions with the female dummy.

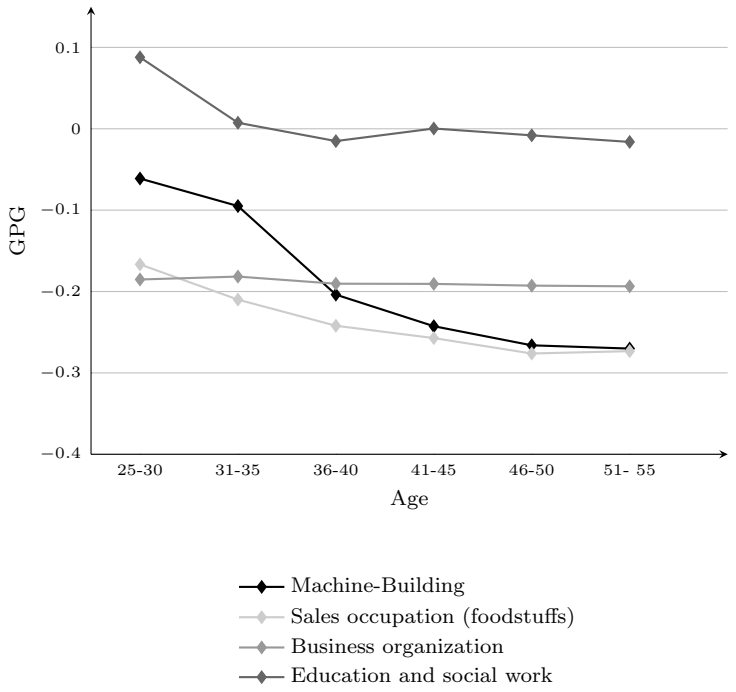
In the male-dominated occupation of Machine-building and -operating, the GPG for the youngest cohort is relatively low (-6 %) and falling to -20 % (36-40), and reaching -27 % in the oldest age groups. The share of women, however, develops in the opposite direction. That is, in the younger cohorts fewer women (in comparison to men) work in this occupation than in the older cohorts. Similar trends can be observed in the full-time sample. These findings could be interpreted as a change in group selection. Hence, fewer women are selecting in this occupations, but those who do, appear to be a very selective group and the extent of positive-selection may be increasing over time.

In addition, in the Sales Occupation, the GPG decreases across age groups. While men aged 25 to 30 earn 17 % more than their female colleagues, men in the oldest age group earn 27 % more than women. Again, we observe an increase in the share of women in the entire sample. However, the share of women in the full-time sample between 31 and 40 years drops. This finding may be related to the fact that women with young children are more likely to work part-time. As a reminder, in the data we observe tenure and age of employees but not whether, and if so, for how long they have worked part-time. This example shows that the adjusted GPG may be overestimated as tenure does not allow for differentiating between part-time and full-time work experience within the firm. Moreover, the data gives no information on duration of employment breaks, which also might lead to an overestimation of the GPG. In addition, we cannot observe whether an employee has been promoted within this occupation, such that she or has more responsibilities, which justifies wage differences. As Blau and DeVaro (2007) shows, men are more likely to be promoted, which may explain some parts of the adjusted GPG.

Comparing Sales Occupation and Education and Social Work, emphasizes the role of linearity of earnings and collective agreements. Employees in Education and Social Work also show a remarkable drop in the share of women in the full-time sample, which, in this occupation, however, is not related to a drop in the GPG. Due to collective agreements, men and women earn the same. Moreover, hierarchical structures are more flat in those occupations, which reduces the impact of unobserved promotions on wage differences.

In Business Organization the GPG is stable across ages, which indicates that even at the beginning of their careers women earn 19 % less than their male colleagues. As men and women might have similar biographies at the beginning of their careers, the wage gap cannot be explained by unobserved career breaks. Rather, this finding highlights the impact of negotiations on the GPG: Wages in Business Organization are often negotiated and women typically are less successful at wage negotiations than men, which results in a larger GPG. The share of women in the entire sample and in the full-time sample shows that women, who are 40 years or older are less likely to work full-time. Based on

these findings, the high degree of non-linearity in earnings in this occupation (see section 2) might also reflect the impact of selection processes into full-time and part-time work. Hence, the positive-selected group of women may work in full-time, while those with lower labor market attachment might select into part-time.

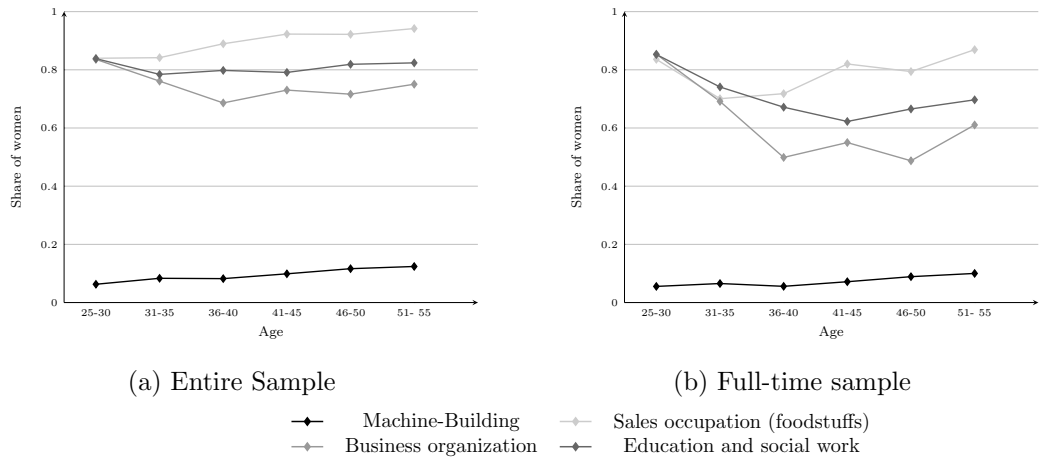


Source: SES 2014; Based on all employees between 25 and 55 years.

Figure 6: Adjusted wage gaps within age groups in selected occupations

In sum, these results indicate that the GPG increases with age but the extent to which the GPG is correlated with age, differs substantially between occupations. This finding may also be linked to missing data. As a reminder, the SES is a cross-sectional data including tenure and age, but not the entire employment biography. Hence, we observe potential but not actual work experience. Further, the data offers no information on preferences and, therefore, does not allow for correcting for selection processes not only in occupations *per se* (e.g. Adda et al., 2017), but also in full- or part-time employment (e.g. Gallego Granados, 2019).

These findings emphasize that the data misses important information that may be related to differences in earnings between men and women such as part-time work experience or career breaks. Therefore, GPGs may be overestimated and the bias may vary across oc-



(a) Entire Sample (b) Full-time sample

Source: SES 2014; Based on all employees between 25 and 55 years.

Figure 7: Share of women in selected occupations

cupations, which might affect the results of equation (1) and to what extent occupational characteristics are correlated with the GPG.

7 Conclusion

The aim of this study is to analyze whether or not, occupational characteristics are able to explain the variation in the GPG in Germany. Based on a two-step approach, I show that the adjusted GPG varies substantially across occupations: The largest gap is observed in the occupation of Actors, Dancers and Athletes (51 %), and in other occupations such as in Event Organization, in contrast, women earn even 6 % more than men.

In the second step, I link the GPG to occupational characteristics such as the share of women or the non-linearity index, that gives the relative difference in hourly wages between persons working more than 40 hours per week and those with less than 25 hours. The results reveal that there are four occupational characteristics that are highly correlated with the GPG: The non-linearity in earnings, the hierarchical composition, the tasks and the ownership of a firm. Hence, the results do not only confirm previous findings from the USA, where selected occupations with non-linear earnings tend to have higher GPGs, but extends them to a more general level. Moreover, there is evidence of a glass ceiling as the GPG increases with the share of persons having supervisory power. In addition, the result supports the findings from the task-based literature, as it shows higher GPGs in occupations with routine tasks. Both, the negative relationship between the non-linearity in earnings and the GPG and higher wage dif-

ferences in routine tasks, emphasize the role of substitution: The more that employees can be substituted with other employees, the lower is the GPG. Finally, collective agreements in public firms result in more equal wages between women and men.

However, the model explains only a quarter of the variance in the GPGs across occupations, thus emphasizing that a major part remains unobserved. Further, information such as actual work-experience in part- and full-time might be correlated with the GPG but cannot be observed in the data. Therefore, the adjusted GPG and, thus, differences between occupations may be overestimated.

References

Adda, Jerome, Christian Dustmann, and Katrien Stevens (2017). "The Career Costs of Children". In: *Journal of Political Economy* 125.2, pp. 293–337.

Allmendinger, Jutta (1989). "Educational systems and labor market outcomes". In: *European Sociological Review* 5.3, pp. 231–250.

Arulampalam, Wiji, Alison L. Booth, and Mark L. Bryan (2007). "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution". In: *ILR Review* 60.2, pp. 163–186.

Babcock, Linda and Sara Leschever (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press.

Bertrand, Marianne (2011). "New perspectives on gender". In: *Handbook of Labor economics* 4, pp. 1543–1590.

Bhalotra, Sonia R. and Manuel Fernández (2018). *The Distribution of the Gender Wage Gap*. DP 11640. IZA.

Black, Sandra E. and Alexandra Spitz-Oener (2010). "Explaining women's success: Technological change and the skill content of women's work". In: *The Review of Economics and Statistics* 92.1, pp. 187–194.

Blau, Francine D and Jed DeVaro (2007). "New evidence on gender differences in promotion rates: An empirical analysis of a sample of new hires". In: *Industrial Relations: A Journal of Economy and Society* 46.3, pp. 511–550.

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent, Trends and Explanations". In: *Journal of Economic Literature* 55.3, pp. 789–865.

Bundesagentur für Arbeit (2011). *Klassifikation der Berufe 2010 – Band 1: Systematischer und alphabetischer Teil mit Erläuterungen*. Nuremberg.

Busch, Anne and Elke Holst (2009). *Glass Ceiling Effect and Earnings: The Gender Pay Gap in Managerial Positions in Germany*. Discussion Paper 905. Berlin: DIW Berlin.

— (2011). *Gender-Specific Occupational Segregation, Glass Ceiling Effects, and Earnings in Managerial Positions: Results of a Fixed Effects Model*. Discussion Paper 1101. Berlin: DIW Berlin.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). "Bargaining, sorting, and the gender wage gap: quantifying the impact of firms on the relative pay of women". In: *Quarterly Journal of Economics* 131.2, pp. 633–686.

Collischon, Matthias. "Is there a Glass Ceiling over Germany?" In: *German Economic Review* 175. forthcoming.

Coudin, Elise, Sophie Maillard, and Maxime Tô (2018). *Family, Firms and the Gender Wage Gap in France*. IFS Working Paper W18/01. Institute for Fiscal Studies.

Croson, Rachel and Uri Gneezy (2009). "Gender Differences in Preferences". In: *Journal of Economic Literature* 47.2, pp. 1–27.

Dengler, Katharina, Britta Matthes, and Wiebke Paulus (2014). *Berufliche Tasks auf dem deutschen Arbeitsmarkt*. FDZ Methodenreport 12/2014. Nuremberg: Institute of Employment Research (IAB).

Finke, Claudia, Florian Dumpert, and Martin Beck (2017). *Verdienstunterschiede zwischen Männern und Frauen*. WISTA 2. Statistisches Bundesamt.

Gallego Granados, Patricia (2019). *The Part-Time Wage Gap across the Wage Distribution*. Discussion Paper 1791. DIW Berlin.

Gallego Granados, Patricia and Katharina Wrohlich (2018). *Gender Pay Gap besonders groß bei niedrigen und hohen Löhnen*. Wochenbericht 10/2018. DIW Berlin.

Gartner, Hermann and Thomas Hinz (2009). "Geschlechtsspezifische Lohnungleichheit in Betrieben, Berufen und Jobzellen (1993–2006)". In: *Berliner Journal für Soziologie* 19.4, pp. 557–575.

Gesundheitsberichterstattung des Bundes (2019). *Bei den Ärztekammern registrierte Ärztinnen und Ärzte mit Gebiets- und Facharztbezeichnung*.

Goldin, Claudia (2014). "A Gender Convergence: Its Last Chapter". In: *American Economic Review* 104.4, pp. 1091–1119.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth (2017). "The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census". In: *American Economic Review: Papers & Proceeding* 107.5, pp. 110–114.

Hausmann, Ann-Christin, Corinna Kleinert, and Kathrin Leuze (2015). "Entwertung von Frauenberufen oder Entwertung von Frauen im Beruf?" In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 67, pp. 217–242.

Hinz, Thomas and Hermann Gartner (2005). "Geschlechtsspezifische Lohnunterschiede in Branchen, Berufen und Betrieben/The Gender Wage Gap within Economic Sectors, Occupations, and Firms". In: *Zeitschrift für Soziologie* 34.1, pp. 22–39.

Hirsch, Boris (2013). "The impact of female managers on the gender pay gap: Evidence from linked employer–employee data for Germany". In: *Economics Letters* 119.3, pp. 348–350.

Kapteyn, Arie and Jelmer Y. Ypma (2007). "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data". In: *Journal of Labor Economics* 25.3, pp. 513–551.

Levanon, Asaf, Paula England, and Paul Allison (2009). "Occupational Feminization and Pay: Assessing Causal Dynamics Using 1950-2000 U.S. Census Data". In: *Social Forces* 88.2, pp. 865–892.

Ludsteck, Johannes (2014). "The impact of segregation and sorting on the gender wage gap - evidence from German linked longitudinal employer-employee data". In: *ILR Review* 67.2, pp. 362–394.

Matthes, Britta, Holger Meiniken, and Petra Neuhauser (2015). *Berufssektoren und Berufssegmente auf Grundlage der KldB 2010*. Methodenbericht. Bundesagentur für Arbeit Statistik.

Murphy, Emily and Daniel Oesch (2016). "The Feminization of Occupations and Change in Wages: A Panel Analysis of Britain, Germany, and Switzerland". In: *Social Forces* 94.3, pp. 1221–1255.

Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?" In: *Quarterly Journal of Economics* 129.3, pp. 1067–1101.

Paulus, Wiebke and Britta Matthes (2013). *Klassifikation der Berufe : Struktur, Codierung und Umsteigeschlüssel*. FDZ Methodenreport 08/2013. Nuremberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).

Spitz-Oener, Alexandra (2006). "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure". In: *Journal of Labor Economics* 24.2, pp. 235–270.

Statistisches Bundesamt (2018a). *Abhängig Erwerbstätige: Deutschland, Jahre, Beschäftigungsumfang, Geschlecht*.

— (2018b). *Unternehmen und Arbeitsstätten 2015*. Fachserie 2. Fachserie 2 Reihe 1.6.1.

— (2018c). *Verdienststrukturerhebung*.

Weeber, Sonja and Dietmar Hobler (2015). *LehramtsstudentInnen nach Studiengängen und Fächern im Wintersemester 2013/14 in Deutschland*. GEW.

Appendix

Classification	Tasks
non-routine analytic	researching, analyzing, evaluating and planning, making plans/ constructions, designing, sketching, working out rules/ prescriptions, and using interpreting rules
non-routine interactive	negotiating, lobbying, coordinating, organizing, teaching or training, selling, buying, advising customers, advertising, entertaining or presenting, and employing or managing personnel
routine cognitive	calculating, bookkeeping, correcting texts/ data, and measuring length/ weight/ temperature
routine manual	operating or controlling machines and equipping machines
non-routine manual	repairing or renovating houses/ apartments/ machines/ vehicles, restoring art/ monuments, and serving or accommodating

Source: Spitz-Oener (2006)

Table 6: Activities within tasks

Table 7: Occupations and occupational segments according to the German Classification of Occupation 2010

Code	**Classification title**
111	Occupations in farming
112	Occupations in animal husbandry
113	Occupations in horsekeeping
114	Occupations in fishing
115	Occupations in animal care
116	Occupations in vini- and viticulture
117	Occupations in forestry, hunting and landscape preservation
121	Occupations in gardening
122	Occupations in floristry
211	Occupations in underground and surface mining and blasting engineering
212	Conditioning and processing of natural stone and minerals, production of building materials
213	Occupations in industrial glass-making and -processing
214	Occupations in industrial ceramic-making and -processing
221	Occupations in plastic- and rubber-making and -processing
222	Occupations in colour coating and varnishing
223	Occupations in wood-working and -processing
231	Technical occupations in paper-making and -processing and packaging

Continued on next page

Table 7 – continued from previous page

Code	Classification title
232	Occupations in technical media design
233	Occupations in photography and photographic technology
234	Occupations in printing technology, print finishing, and book binding
241	Occupations in metal-making
242	Occupations in metalworking
243	Occupations in treatment of metal surfaces
244	Occupations in metal constructing and welding
245	Occupations in precision mechanics and tool making
251	Occupations in machine-building and -operating
252	Technical occupations in the automotive, aeronautic, aerospace and ship building industries
261	Occupations in mechatronics, automation and control technology
262	Technical occupations in energy technologies
263	Occupations in electrical engineering
271	Occupations in technical research and development
272	Draftspersons, technical designers, and model makers
273	Technical occupations in production planning and scheduling
281	Occupations in textile making
282	Occupations in the production of clothing and other textile products
283	Occupations in leather- and fur-making and -processing
291	Occupations in beverage production
292	Occupations in the production of foodstuffs, confectionery and tobacco products
293	Cooking occupations
311	Occupations in construction scheduling and supervision, and architecture
312	Occupations in surveying and cartography
321	Occupations in building construction
322	Occupations in civil engineering
331	Floor layers
332	Painters and varnishers, plasterers, occupations in the waterproofing of buildings, preservation of structures and wooden building components
333	Occupations in the interior construction and dry walling, insulation, carpentry, glazing, roller shutter and jalousie installation
341	Occupations in building services engineering
342	Occupations in plumping, sanitation, heating, ventilating, and air conditioning
343	Occupations in building services and waste disposal
411	Occupations in mathematics and statistics
412	Occupations in biology
413	Occupations in chemistry
414	Occupations in physics
421	Occupations in geology, geography and meteorology
422	Occupations in environmental protection engineering
423	Occupations in environmental protection management and environmental protection consulting
431	Occupations in computer science

Continued on next page

Table 7 – continued from previous page

Code	Classification title
432	Occupations in IT-system-analysis, IT-application-consulting and IT-sales
433	Occupations in IT-network engineering, IT-coordination, IT-administration and IT-organization
434	Occupations in software development and programming
511	Technical occupations in railway, aircraft and ship operation
512	Occupations in the inspection and maintenance of traffic infrastructure
513	Occupations in warehousing and logistics, in postal and other delivery services, and in cargo handling
514	Service occupations in passenger traffic
515	Occupations in traffic surveillance and control
516	Management assistants in transport and logistics
521	Driver of vehicles in road traffic
522	Drivers of vehicles in railway traffic
523	Aircraft pilots
524	Ship's officers and masters
525	Drivers and operators of construction and transportation vehicles and equipment
531	Occupations in physical security, personal protection, fire protection and workplace safety
532	Occupations in police and criminal investigation, jurisdiction and the penal institution
533	Occupations in occupational health and safety administration, public health authority, and disinfection
541	Occupations in cleaning services
611	Occupations in purchasing and sales
612	Trading occupations
613	Occupations in real estate and facility management
621	Sales occupations in retail trade (without product specialization)
622	Sales occupations (retail trade) selling clothing, electronic devices, furniture, motor vehicles and other durables
623	Sales occupations (retail) selling foodstuffs
624	Sales occupations (retail) selling drugstore products, pharmaceuticals, medical supplies and healthcare goods
625	Sales occupations (retail) selling books, art, antiques, musical instruments, recordings or sheet music
631	Occupations in tourism and the sports (and fitness) industry
632	Occupations in hotels
633	Gastronomy occupations
634	Occupations in event organization and management
711	Managing directors and executive board members
712	Legislators and senior officials of special interest organizations
713	Occupations in business organization and strategy
714	Office clerks and secretaries
715	Occupations in human resources management and personnel service
721	Occupations in insurance and financial services
722	Occupations in accounting, controlling and auditing

Continued on next page

Table 7 – continued from previous page

Code	Classification title
723	Occupations in tax consultancy
731	Occupations in legal services, jurisdiction, and other officers of the court
732	Occupations in public administration
733	Occupations in media, documentation and information services
811	Doctors' receptionists and assistants
812	Laboratory occupations in medicine
813	Occupations in nursing, emergency medical services and obstetrics
814	Occupations in human medicine and dentistry
815	Occupations in veterinary medicine and non-medical animal health practitioners
816	Occupations in psychology and non-medical psychotherapy
817	Occupations in non-medical therapy and alternative medicine
818	Occupations in pharmacy
821	Occupations in geriatric care
822	Occupations providing nutritional advice or health counselling, and occupations in wellness
823	Occupations in body care
824	Occupations in funeral services
825	Technical occupations in medicine, orthopaedic and rehabilitation
831	Occupations in education and social work, and pedagogic specialists in social care work
832	Occupations in housekeeping and consumer counselling
833	Occupations in theology and church community work
841	Teachers in schools of general education
842	Teachers for occupation-specific subjects at vocational schools and in-company instructors in vocational training
843	Teachers and researcher at universities and colleges
844	Teachers at educational institutions other than schools (except driving, flying and sports instructors)
845	Driving, flying and sports instructors at educational institutions other than schools
911	Occupations in philology
912	Occupations in the humanities
913	Occupations in the social sciences
914	Occupations in economics
921	Occupations in advertising and marketing
922	Occupations in public relations
923	Occupations in publishing and media management
924	Occupations in editorial work and journalism
931	Occupations in product and industrial design
932	Occupations in interior design, visual marketing, and interior decoration
933	Occupations in artisan craftwork and fine arts
934	Artisans designing ceramics and glassware
935	Artisans working with metal
936	Occupations in musical instrument making
941	Musicians, singers and conductors

Continued on next page

Table 7 – continued from previous page

Code	Classification title
942	Actors, dancers, athletes and related occupations
943	Presenters and entertainers
944	Occupations in theatre, film and television productions
945	Occupations in event technology, cinematography, and sound engineering
946	Occupations in stage, costume and prop design
947	Technical and management occupations in museums and exhibitions
011*	Commissioned officers
012*	Senior non-commissioned officers and higher
013*	Junior non-commissioned officers
014*	Armed forces personnel in other ranks

Source: Bundesagentur für Arbeit (2011) and Matthes, Meiniken, and Neuhauser (2015);
Note: Occupations with * are not included in the analysis as they do not offer task information
Occupational sectors: S1: Production, S2: Person-related services, S3: Business administration and business-related services, S4: IT and science related services, S5: Other economic services

	(I)	(II)
Age	0.031	0.034
	(0.001)	(0.001)
Age^2	−0.000	−0.000
	(0.000)	(0.000)
Tenure	0.018	0.014
	(0.000)	(0.000)
$Tenure^2$	−0.000	−0.000
	(0.000)	(0.000)
Permanent contract	0.097	0.118
	(0.002)	(0.002)
Former GDR	−0.185	−0.205
	(0.001)	(0.001)
Leadership Position	0.260	0.245
	(0.002)	(0.002)
Size of Establishment (ref: 50-499 employees)		
1-9 employees	−0.082	−0.085
	(0.005)	(0.006)
10-49 employees	−0.040	−0.051
	(0.003)	(0.003)
>500 employees	0.091	0.100
	(0.002)	(0.002)
Size of company (ref: 50-249 employees)		
1-9 employees	−0.124	−0.092
	(0.005)	(0.006)
10-49 employees	−0.052	−0.035
	(0.003)	(0.003)
250-1000 employees	0.053	0.050
	(0.002)	(0.002)
> 1000 employees	0.126	0.112
	(0.002)	(0.002)
Urban region	0.060	0.064
	(0.001)	(0.001)
Education (ref: No A-Level. VT)		
No A-Level. No VT	−0.081	−0.092
	(0.002)	(0.002)
A-Level. No VT	−0.010	0.089
	(0.003)	(0.004)
A-Level. VT	0.085	0.095
	(0.002)	(0.002)
Polytechnical degree	0.160	0.196
	(0.003)	(0.004)
University degree	0.333	0.345
	(0.002)	(0.002)
N	434,821	315,289
R^2	0.615	0.612

Source: SES 2014; *Note*: Not shown: Constant, occupational fixed effects and female-occupational fixed effects.; (I): Entire sample (min. working hours 9 hs/week), (II): Full-time sample (min. working hours 35 hs/week); Standard errors presented in parentheses; All coefficients are statistically significant.

Table 8: Results: First step