DE GRUYTER OLDENBOURG Journal of Economics and Statistics 2023; 243(2): 183–195

Data Observer

Matthias Collischon*

Identifying Supervisory or Managerial Status in German Administrative Records

https://doi.org/10.1515/jbnst-2022-0035
Received May 23, 2022; accepted May 24, 2022

Abstract: Information on individuals holding managerial or supervisory positions within establishments is important for various aspects of labour market research. However, identifying managers or supervisors in German administrative records is not straightforward. This paper uses survey information from the Panel Study Labour Market and Social Security (PASS) to predict managerial or supervisory tasks in administrative records that can be used to enhance the identification of managers and supervisors in the Sample of Integrated Labour Market Biographies (SIAB). Furthermore, I provide an applied example in which I calculate gender differences in the probability to hold a managerial position.

Keywords: managers, PASS (Panel Study Labour Market and Social Security), SIAB (Sample of Integrated Labour Market Biographies), imputation

JEL Classification: C53, J16

1 Introduction

German register data have become popular in various fields, e.g. economics (Card et al. 2013) and sociology (Huffman et al. 2017) due to its large sample size and daily information on employment characteristics. However, not every variable that is included in the data is reliable as they are not essential for social security contributions. For example, Fitzenberger et al. (2005) show that the measure for education in the data, as it is reported, is sometimes inconsistent and contains many missing values. The same also holds true for several other information in the data, including the measure for occupation and, specifically, whether an individual holds managerial or supervisory duties.

***Corresponding author: Matthias Collischon**, Institute for Employment Research (IAB), Regensburger Straße 100, 90478 Nürnberg, Germany, E-mail: matthias.collischon2@iab.de

The access to managerial positions and the consequences of managers' characteristics for various processes in the labour market are a growing field in the social sciences literature. One specific example is the glass ceiling, i.e. the reduced chance to obtain a managerial position, for women in the labour market. Maume (1999), for example, shows that women face barriers in the access to managerial positions compared to men. To investigate such research questions with the IAB data, it is thus important to precisely identify managerial positions.

This paper proposes a procedure to improve the identification of individuals with supervisory or managerial duties in the employee history (*Beschäftigtenhistorik*, BeH) data. It uses survey information from the Panel Study Labour Market and Social Security (PASS) that can be linked with administrative records to predict the probability of being a manager using only information that is also available in the administrative records. As supplementary material, I provide data and syntax to use my results to predict managers in the Sample of Integrated Labour Market Biographies (SIAB) data. This supplemental material was created using Stata 16.

Specifically, the procedure I implement works by estimating a regression in which the information on managerial duties from the survey data is regressed on information available in administrative data. This provides coefficients for the variables included in the administrative data to predict managerial status. Thus, I next use the estimated vector of coefficients in the full sample of administrative information in the SIAB to predict managerial status in this context. I use the predicted probabilities and various thresholds (e.g. 70%) to predict managerial status as a binary variable. Next, I use these predictions to estimate the gender gap in holding managerial positions in the PASS and SIAB data.

2 Identifying Managers in German Administrative Data

2.1 The Status Quo: Using the Classification of Occupations

Identifying managers in German administrative data is theoretically possible by using the information on the classification of occupations (*Klassifikation der Berufe*, KldB; Statistisches Bundesamt 2020). If the fourth digit of the KldB code is a "9", this indicates that the employee has managerial or supervisory duties. Furthermore, the KldB Code "7110" indicates being an executive. Theoretically, using these codes should suffice in identifying managers.

However, in practice, employers have no incentive to update this information if a promotion happens. This means, for example, if an employee starts working at an establishment of a given firm without managerial duties and is promoted to a managerial position, it could be the case that she simply keeps her original KldB information in the administrative data and the employer only updates information on pay and the contract duration. Thus, we would falsely identify her as a non-manager.

Furthermore, not all employees with managerial duties are necessarily classified via a “9” in the KldB-codes as managers, because the classification does not allow for it (as some occupations do not contain this category), as also noted by Paulus and Matthes (2013).

Thus, it is likely the case that any researcher using only the KldB-codes to identify employees with managerial duties underestimates the true proportion of these and that the definition for managers or employees with supervisory positions according to the occupational classification is far stricter than what survey information would suggest.

2.2 Enhancing Managerial Information using PASS-Data

The main part of this paper uses information from the Panel Study Labour Market and Social Security (PASS) (Trappmann et al. 2019), linked with administrative records (PASS-ADIAB) (Antoni and Bethmann 2019) to use survey information on managerial duties to identify managers in administrative records. The PASS is a panel study that consists of a sample of German households that oversamples welfare benefits and started in 2006. Yearly, it surveys around 10,000 households with around 15,000 individuals. Respondents are asked to allow for consent record linkage of their survey information to administrative records; around 80% of respondents give their consent (Antoni and Bethmann 2019). This unique setting allows me to use information from administrative records to predict variables that are part of the survey questionnaire. Even though benefit recipients are over-sampled, the subsample of individuals in employment is still large enough for my analysis.

Since 2011, individuals are asked whether they supervise or are authorized to issue instructions to other employees. I use this questionnaire item to predict supervisory or managerial status in administrative records. This is achieved by regressing managerial status as it is reported in the survey on a set of variables from the administrative data and then uses the coefficients from this regression to estimate the propensity of being a manager in administrative records.

To prevent risks from overfitting, I begin by randomly splitting the PASS-data that is linked to administrative records into a train and a test sample (50% of the full sample, respectively). Table 1 shows sample descriptives of the relevant variables for both samples. As can be seen, the number of individuals that report any supervisory or managerial duties in the survey (36%) exceeds the number of individuals that are identified as managers by using the occupational classification (3%) by far. As a sidenote, 36% of employees having managerial or supervisory duties may seem large, but data from the Mikrocensus 2016 reports 24% of working individuals with supervisory or managerial duties (DeStatis 2020, Item EF120). This indeed suggests that a large number of employees with managerial or supervisory duties are not reported correctly in the administrative records or that administrative information rely on a far stricter definition of managerial and supervisory positions. Either way, especially to ensure comparability of estimations based on administrative records with survey information, it can be useful to build a variable that resembles the survey information more closely.

In the next step, I estimate a regression model via probit and logit to predict the probability of being a manager in the train data. I then use the predictions from this model in the test data to assess the quality of the procedure and compare the predictions to the managerial information in the survey. Specifically, I estimate the following model:

Table 1: Sample descriptives for train and test data.

Variable	Train Data		Test Data	
	Mean	SD	Mean	SD
Manager (PASS) (0/1)	0.36	0.48	0.36	0.48
Manager (KldB) (0/1)	0.03	0.18	0.03	0.18
Daily pay (Euros)	73.92	43.75	73.35	43.43
Age	43.35	10.83	43.34	10.90
Female (0/1)	0.51	0.50	0.52	0.50
Current year	2012.19	3.14	2012.19	3.14
Labour market experience (years)	16.01	9.36	15.99	9.38
Job tenure (years)	4.63	5.93	4.62	6.00
Education: No vocational training	0.06	0.24	0.07	0.25
Education: Vocational training	0.62	0.48	0.62	0.49
Education: Upper secondary	0.03	0.17	0.03	0.17
Education: Upper secondary + voc training	0.12	0.32	0.12	0.33
Education: University of applied sciences	0.04	0.19	0.04	0.19
Education: University	0.13	0.34	0.13	0.34
Observations	20,337		20,736	

Source: PASS Survey, own calculations.

$$\Pr(Manager = 1)_{it} = \beta_0 + \beta_1 Kldb'_{it} + \beta_2 X'_{it}$$

Thus, I use information on the classification of occupations as well as covariates to estimate the probability of being a manager. I use three different variations of the occupational classification variable:

i) A simple indicator (0/1) for being a manager according to the classification
ii) A set of indicator variables for the 3-digit-classification
iii) A set of indicator variables for the 4-digit-classification

X' contains daily pay (cubic polynomial), labour market experience and job tenure (both with squared polynomials), gender, education dummies and age and survey year (dummies). The model is estimated via a logit and a probit regression for robustness.

Table 2 shows the coefficients from the estimations (the KldB-category coefficients are not shown for brevity), with columns 1–3 showing the logit estimation results and columns 4–6 showing the probit results, respectively.[1] As can be seen, pay correlates positively with managerial status, as does education, gender and tenure. The picture with regards to experience is mixed, suggesting an inverted u-shaped connection.

Next, I use the estimation coefficients to predict the probability to hold a managerial or supervisory position. To assess the quality of the prediction overall, I use ROC-curves. The results are displayed in Figure 1. As can be seen, the predictions perform relatively well in terms of sensitity, even with a low rate of false positives, regardless of using a probit or logit estimation for prediction; there is hardly any difference between logit or probit. Furthermore, the curves from the estimations using the 3- and 4-digit-KldB show negligible differences, while the estimations using only the binary managerial status indicator seem to perform substantially worse. Overall, the implications from these curves are that predicting managerial or supervisory status works fairly well, regardless of the method used.

In the next step, I use the predicted probabilities from this regression to predict a binary variable of managerial status in the PASS data. Because the prediction performs relatively well out of sample, I now use the pooled PASS sample for the prediction to increase statistical power in the estimations that follow. Due to the negligible differences between the logit- and probit predictions, I now rely on the logit estimation to predict managerial status and offer several variants of identification, each providing a dummy for being a manager when

1 Table A1 shows the corresponding marginal effects for the logit- and probit-coefficient estimates.

Table 2: Comparing various predictions against survey information on managerial status.

	(1) Binary dummy	(2) Logit 3-digit-KldB	(3) 4-digit-KldB	(4) Binary dummy	(5) Probit 3-digit-KldB	(6) 4-digit-KldB
Daily pay	0.016[c]	0.012[b]	0.012[b]	0.009[c]	0.005[a]	0.006[b]
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Daily pay squared	−0.000	0.000	0.000	−0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Daily pay cubic	0.000	−0.000	−0.000	0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No vocational training	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Vocational training	0.341[c]	0.278[c]	0.305[c]	0.200[c]	0.164[c]	0.179[c]
	(0.078)	(0.082)	(0.084)	(0.046)	(0.047)	(0.048)
Upper Secondary	0.275	0.211	0.264[a]	0.167[a]	0.137	0.166[a]
	(0.122)	(0.127)	(0.130)	(0.072)	(0.074)	(0.075)
Upper secondary + voc training	0.384[c]	0.406[c]	0.468[c]	0.229[c]	0.244[c]	0.276[c]
	(0.089)	(0.094)	(0.097)	(0.052)	(0.055)	(0.056)
University of applied sciences	0.364[b]	0.339[b]	0.365[b]	0.216[b]	0.200[b]	0.215[b]
	(0.112)	(0.118)	(0.122)	(0.066)	(0.070)	(0.072)
University	0.345[c]	0.293[b]	0.326[b]	0.200[c]	0.173[b]	0.189[b]
	(0.091)	(0.099)	(0.104)	(0.053)	(0.058)	(0.061)
Female (0/1)	−0.259[c]	−0.434[c]	−0.409[c]	−0.160[c]	−0.259[c]	−0.245[c]
	(0.034)	(0.044)	(0.046)	(0.021)	(0.026)	(0.027)
Experience	0.048[c]	0.046[c]	0.048[c]	0.029[c]	0.028[c]	0.028[c]
	(0.008)	(0.008)	(0.009)	(0.005)	(0.005)	(0.005)
Experience squared	−0.001[c]	−0.001[b]	−0.001[b]	−0.000[c]	−0.000[b]	−0.000[b]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Job tenure	0.044[c]	0.045[c]	0.048[c]	0.027[c]	0.027[c]	0.029[c]
	(0.007)	(0.007)	(0.008)	(0.004)	(0.004)	(0.005)
Job tenure squared	−0.001[b]	−0.001[b]	−0.001[b]	−0.001[b]	−0.001[b]	−0.001[b]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	−2.539[c]	0.104	0.009	−1.496[c]	0.174	0.140
	(0.444)	(1.196)	(1.203)	(0.267)	(0.694)	(0.697)
Pseudo R2	0.092	0.128	0.151	0.091	0.128	0.150
N	19,492	19,456	18,966	19,492	19,456	18,966

Source: PASS Survey, own calculations. Standard errors in parentheses. Dependent variable: managerial status (PASS survey item). [a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.001$.

i) the prediction from the binary indicator predicts an above 80% chance to be a manager (i.e. the predicted probability in the logit estimation is larger or equal to 0.8)

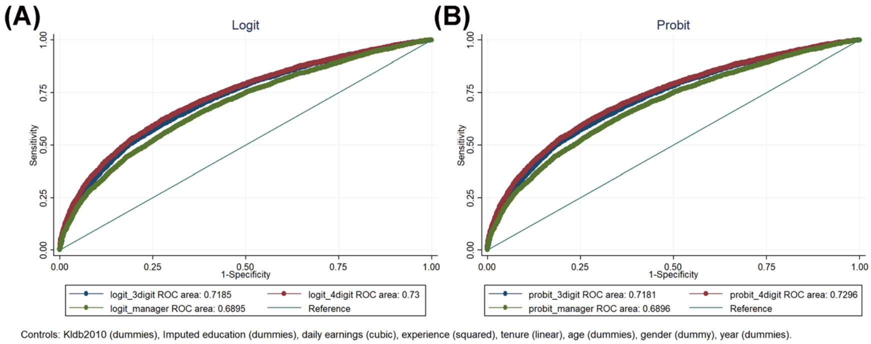

Figure 1: ROC-curves for logit and probit estimation.

ii) the prediction from the model including the 3-digit-KldB predicts an above 80% chance to be a manager
iii) the prediction from the model including the 4-digit-KldB predicts an above 80% chance to be a manager
iv) the prediction from the model including the 3-digit-KldB predicts an above 70% chance to be a manager

These varying definitions test both the sensitivity as well as the quality of the prediction. Choosing the thresholds at 70% or 80% implies that the predictions have to be far more certain than flipping a coin, but not too restrictive. At this point, I would also like to note that the supporting information for this article contain the parameter estimates to predict the likelihood of holding a managerial position for all empirical models. Thus, researchers do not need to rely on my definitions of managerial status, but can build their own classifications. These definitions are, however, several benchmarks that can provide guidance in applied research.

Table 3: Comparing various predictions against survey information on managerial status.

(A) Simple prediction (*N* = 39,388)	**Manager (PASS)**		**(B) 3-digit-KldB (*N* = 39,350)**	**Manager (PASS)**	
	No	Yes		No	Yes
No	24,936	112	No	24,648	371
Yes	13,595	745	Yes	12,678	1653
(C) 3-digit-KldB, 70% threshold (*N* = 39,350)	Manager (PASS)		(D) 4-digit-KldB (*N* = 38,643)	Manager (PASS)	
	No	Yes		No	Yes
No	24,454	565	No	24,428	187
Yes	12,284	2047	Yes	12,634	1394

Source: PASS Survey, own calculations.

Table 3 provides cross tabulations of the various definitions of being a manager versus the survey question. As can be seen, the risk of false positives in any case is relatively low, and never higher than 1.5% (with the 3-digit-KldB classification and a 70% threshold for prediction). However, in either case, the identification provides false-negatives in around one third of cases. Nevertheless, either prediction identifies more managers in the data than simply using the original occupation classification that would only identify 1300 observations as managers, whereas the procedure using the 3-digit-KldB and the 70% threshold predicts around 2660 observations for managers.

3 Empirical Example: Gender Differences in Holding Managerial or Supervisory Jobs

3.1 How do the Predictions Perform in Regressions Compared to Survey Information in the PASS?

In the next step, I compare the results of the survey information to the various predictions using the example of the glass ceiling in holding managerial positions for females. Thus, I use managerial status as an outcome and estimate linear probability estimations[2] of the following model:

$$manager_{it} = \delta_0 + \delta_1 female_{it} + \delta_2 C'_{it} + \epsilon_{it}$$

Where *manager* is the binary outcome of being a manager, *female* is an indicator variable for the individual's gender, and C is a set of control variables available in the administrative data, namely, experience and tenure (both with squared polynomials), current year (dummies), educational degree and age dummies. I want to make clear that this is only a relatively simple empirical approach to the problem. In a more formal and sophisticated analysis, one would not only investigate mean differences in the propensity to hold a supervisory or managerial position, but also differential returns for various covariates by gender. I nonetheless consider this simple example as sufficient to gain a grasp on how the predicted variable performs compared to the survey information.

Table 4 shows the results of the regressions as well as the mean of the dependent variable for males. As can be seen using the survey information, women

2 I use linear probability models instead of logit- or probit-estimations to directly estimate the average marginal effect (AME). Additional analyses (not shown) show that the results presented here hardly differ from the AMEs obtained from probit, logit or clog–log estimations.

Table 4: Regression results.

	(1) Manager (survey)	(2) Manager (KldB)	(3) Manager (Prediction w/simple indicator)	(4) Manager (Prediction w/3-digit-KldB)	(5) Manager (Prediction w/4-digit-KldB)	(6) Manager (Prediction w/3-digit-KldB, 70% threshold)
Female (0/1)	−0.111[c]	−0.034[c]	−0.034[c]	−0.057[c]	−0.063[c]	−0.104[c]
	(0.010)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)
Male mean	0.428	0.0544	0.0546	0.0846	0.0942	0.148
Observations	41,073	39,393	39,388	39,350	38,830	39,350

Source: PASS Survey, own calculations. Cluster-robust standard errors in parentheses. [a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.001$.

have 11 percentage points lower chance to hold a managerial position; 42.8% of males report holding supervisory or managerial duties. Using just the KldB classification at face value, the coefficient drops to 3.4 percentage points, but the comparison now is a far lower male baseline of just 5.4%. Of the different estimation procedures, the 3-digit-KldB prediction with the 70% threshold values seems to perform closest to the survey information with regards to the coefficient estimate. However, in this case, the share of males reporting managerial duties is still far lower compared to the survey question. Nevertheless, I take this result as evidence that the 3-digit-KldB prediction performs sufficiently well to create a variable for managerial status that is more in line with the survey information than just using the occupational code.

3.2 The Glass Ceiling in the SIAB

In the next step, I use the vector of coefficients obtained in the prediction with the PASS data to predict managerial status in the SIAB and estimate glass ceilings. This can easily be achieved by loading the estimation results into Stata when the SIAB data is prepared in the same way as the PASS-ADIAB data with regards to variable names. The syntax to replicate this exercise as well as the set of coefficients necessary to replicate these findings is available as supplementary material to this paper online. In this case, I only use the SIAB for the period the PASS is available, i.e. from 2007 to 2017.

Table 5 shows the sample descriptives for the SIAB, including the predictions of managerial status and all covariates also shown in the PASS-data. As can be seen, the share of managers identified through the various prediction procedures

Table 5: Sample descriptives for SIAB data.

Variable	Mean	SD
Manager (KldB only)	0.03	0.18
Manager simple code (prediction)	0.03	0.18
Manager 3-digit-KldB (prediction)	0.05	0.21
Manager 3-digit KldB; 70% threshold (prediction)	0.08	0.28
Manager 4-digit KldB (prediction)	0.05	0.23
Daily pay (Euros)	63.71	55.74
Age	39.85	13.41
Female (0/1)	0.48	0.50
Current year	2012.10	3.16
Labour market experience (years)	13.43	10.09
Job tenure (years)	3.66	5.85
Education: No vocational training	0.19	0.39
Education: Vocational training	0.54	0.50
Education: Upper secondary	0.05	0.21
Education: Upper secondary + voc training	0.09	0.29
Education: University of applied sciences	0.03	0.16
Education: University	0.10	0.31
Observations	13,780,006	

Source: SIAB7517, own calculations.

varies, but is substantially larger than just the KldB code for managers, with the exception of just using the simple indicator for prediction. As previously mentioned this can either indicate that managerial status in the administrative records is measured with noise, or relies on a stricter definition as the survey information. Due to the different data source and procedures to obtain the sample, the SIAB consists of individuals that are more attached to the labour market, whereas the PASS oversamples individuals in long-term unemployment. This distinction is important to keep in mind, but should not fundamentally affect the analysis, as the prediction and imputation only rely on employed individuals.

In the next step, I regress managerial status on gender and a set of covariates. Table 6 shows the results. Again, like in the PASS, the classification of the KldB provides a far smaller absolute gender gap compared to the broader definitions of managerial status using the predictions from the logit-regression. The findings are comparable to the PASS-estimations which suggest that, also in this data, the 3-digit prediction with the 70% threshold seems to provide a broader definition of managerial or supervisory status, arguably more in line with the survey item. Due to the large sample size, all coefficients differ significantly from each other – even

Table 6: Regression results with SIAB data.

	(1) Manager (KLDB)	(2) Manager (Prediction w/simple indicator)	(3) Manager (Prediction w/3-digit-KldB)	(4) Manager (Prediction w/4-digit-KldB)	(5) Manager (Prediction w/3-digit-KldB, 70% threshold)
Female (0/1)	-0.024^{c}	-0.025^{c}	-0.043^{c}	-0.048^{c}	-0.085^{c}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male mean	0.045	0.047	0.072	0.081	0.130
Observations	13,116,091	12,980,802	12,958,975	12,775,566	12,958,975

Source: SIAB7517, own calculations. Cluster-robust standard errors in parentheses. $^{a}p < 0.05$, $^{b}p < 0.01$, $^{c}p < 0.001$.

though using just the KLDB code in column 1 and the prediction with this indicator in column 2 provides arguably the same coefficient in terms of magnitude.

Thus, although there is no survey information available in the SIAB, I would recommend using the definition in column 5 for analyses, as it seems to be most consistent with the survey item from the PASS.

4 Conclusion

In the German administrative records, identifying individuals with managerial or supervisory positions is not straightforward. In this report, I use survey information to identify individuals with managerial responsibilities in the employment records and provide the corresponding programs to replicate my findings. The results show that using my procedure to enhance the information in administrative records provides a picture that is more fitting to survey information on the share of individuals with managerial duties compared to solely using administrative information.

Furthermore, I use the glass ceiling for women as an applied example for using this information in practice in application to the SIAB data. Like in the PASS-ADIAB, the procedure encompasses a more far-reaching definition of supervisory or managerial positions that includes more employees then only using the KldB-code and thus seems to be more comparable to the survey data.

Acknowledgments: I thank Mark Trappmann for his valuable feedback.
Author contributions: Not applicable as the manuscript is single-authored.
Research funding: This research received no targeted funding.

Availability of data and material: The PASS and SIAB data for replication of this study are available from the research data centre of the IAB: https://fdz.iab.de/. The syntax to replicate this paper's findings is provided as supplementary material and are available online.

Competing interests: I have no competing interests to declare.

Appendix A

Table A1: Average marginal effects corresponding to Table 2.

	(1) Logit Manager code	(2) Logit 3-digit	(3) Logit 4-digit	(4) Probit Manager code	(5) Probit 3-digit	(6) Probit 4-digit
Daily pay (Euros)	0.002[c]	0.002[c]	0.002[c]	0.002[c]	0.002[c]	0.002[c]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No vocational training	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Vocational training	0.067[c]	0.052[c]	0.055[c]	0.065[c]	0.052[c]	0.055[c]
	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.014)
Upper secondary	0.053[a]	0.039	0.047[a]	0.054[a]	0.043	0.051[a]
	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)
Upper secondary + voc training	0.076[c]	0.077[c]	0.086[c]	0.075[c]	0.078[c]	0.086[c]
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
University of applied sciences	0.072[b]	0.064[b]	0.067[b]	0.071[b]	0.064[b]	0.066[b]
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
University	0.068[c]	0.055[b]	0.059[b]	0.066[c]	0.055[b]	0.058[b]
	(0.018)	(0.018)	(0.019)	(0.017)	(0.018)	(0.018)
Female	−0.053[c]	−0.086[c]	−0.078[c]	−0.055[c]	−0.085[c]	−0.078[c]
	(0.007)	(0.009)	(0.009)	(0.007)	(0.009)	(0.009)
Labour market experience (years)	0.005[c]	0.005[c]	0.005[c]	0.005[c]	0.005[c]	0.005[c]
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Job tenure (years)	0.007[c]	0.007[c]	0.007[c]	0.007[c]	0.007[c]	0.007[c]
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	19,492	19,456	18,966	19,492	19,456	18,966

Source: PASS Survey, own calculations. Standard errors in parentheses. [a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.001$.

References

Antoni, M. and Bethmann, A. (2019). PASS-ADIAB – linked survey and administrative data for research on unemployment and poverty. Jahrb. Natl. Stat. 239: 747–756.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west German wage inequality. Q. J. Econ. 128: 967–1015.

DeStatis (2020). *Klassifikation der Berufe 2010 – überarbeitete Fassung. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen*. Wiesbaden: Statistisches Bundesamt.

Fitzenberger, Bernd, Osikominu, Aderonke, and Völter, Robert (2005). *Imputation rules to improve the education variable in the IAB employment subsample. (FDZ-Methodenreport, 03/2005 (en)), Nürnberg*. Institute for Employment Research, Nürnberg.

Huffman, M.L., King, J., and Reichelt, M. (2017). Equality for whom? Organizational policies and the gender gap across the German earnings distribution. ILR Rev. 70: 16–41.

Maume, D.J. (1999). Glass ceilings and glass escalators: occupational segregation and race and sex differences in managerial promotions. Work Occup. 26: 483–509.

Paulus, W. and Matthes, B. (2013). *Klassifikation der Berufe * Struktur, Codierung und Umsteigeschlüssel. (FDZ-Methodenreport, 08/2013 (de))*. Institute for Employment Research (IAB), Nürnberg.

Statistisches Bundesamt (2020). *Datenhandbuchzum Mikrozensus Scientific Use File 2016*. DeStatis, Wiesbaden.

Trappmann, M., Bähr, S., Beste, J., Eberl, A., Frodermann, C., Gundert, S., Schwarz, S., Teichler, N., Unger, S., and Wenzig, C. (2019). Data resource profile: panel study labour market and social security (PASS). Int. J. Epidemiol. 48: 1411–1411g.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/jbnst-2022-0035).