Brunow et al. J Labour Market Res (2022) 56:16
https://doi.org/10.1186/s12651-022-00322-7

Journal for Labour Market Research

ORIGINAL ARTICLE

Open Access

Labor market tightness and individual wage growth: evidence from Germany

Stephan Brunow[1], Stefanie Lösch[2] and Ostap Okhrin[2*]

Abstract

It is often stated that certain occupations in Germany, because of "*Demographic Change*", are dwindling, implying a labor shortage. We investigate the 10-year wage growth of young employees entering the labor market in different occupations. Our findings suggest that regional labor market tightness in occupational fields significantly explains wage growth. Individuals who start their careers in a tighter labor market enjoy higher wage growth than workers in more relaxed labor markets. We identify some occupational fields where the effect is especially strong, such as several engineering groups, IT occupations, technicians, and some commercial occupations. Interestingly, health-care occupations reveal a reverse relation.

Keywords: Labor shortages, Occupation-specific wage curve, Demographic change, Wage growth

JEL: J31 (Wage Level and Structure, Wage Differentials), J44 (Professional Labor Markets, Occupational Licensing), R23 (Regional labor markets)

1 Introduction

Technological progress affects many countries economically and socially and shapes the labor market pattern. Acemoglu and Restrepo (2018) show that inequality rises with automation, digitalization, and new tasks: Robots and machines replace routine tasks performed by low-skilled workers. This potentially leads to higher unemployment and stagnant wages for low-skilled labor. Additionally, digitalization and, more generally, technological progress have created new complex tasks for the control, maintenance, and extension of (digitalization) techniques, requiring workers with relevant specialized knowledge in advanced subjects, natural sciences, and IT skills (often referred to as *STEM* for Science, Technology, Engineering, and Mathematics). Moreover, these technologically complex production processes potentially raise the wages of STEM workers. For instance, mechatronics is an excellent example of the change from "pure" mechanics to the combination of mechanics, electronics, IT, and AI techniques. These effects accelerate wage inequality between different groups of workers. To learn more about *Skilled-Bias Technological Change*, see Goldin and Katz (2009) or Acemoglu and Autor (2011), to name a few. According to Autor et al. (2003), it might not (yet) be possible to replace workers with machines for non-routine tasks, such as housekeeping and hotel-related jobs, and personal care, which mainly belong to the low-paid categories. These tasks might be a reason for job polarization at the margins of the wage distribution: low-paid non-routine tasks are still in demand even though high-paid technologically demanding jobs are becoming increasingly important, while machines can increasingly eliminate middle-wage jobs (for the United States: Autor and Dorn 2009; Autor et al. 2006, 2008; the United Kingdom: Goos et al. 2014; and West-Germany: Spitz-Oener 2006; Dustmann et al. 2014). Thus, changing labor demand might cause different wage effects depending on the relative position of individuals within the wage distribution. In countries such as Germany, there is a strong belief that specialists in certain occupations,

*Correspondence: ostap.okhrin@tu-dresden.de

[2] Chair of Statistics and Econometrics, Esp. in Transportation, Faculty of Transportation "Friedrich List", Technische Universität Dresden, Dresden, Germany
Full list of author information is available at the end of the article

especially in STEM, have been *missing* for years (Federal Employment Agency Germany 2017).

Aside from technological progress, the so-called *Demographic Change* made its first appearance in Germany: The (relative) number of young people graduating from college has become smaller since the mid-1990s. Therefore, companies compete aggressively for these youngsters to fill their apprenticeship workplaces (Statistisches Bundesamt 2017).

Is there a shortage of young labor in specific occupations, such as STEM, where labor demand is high, but the supply of (new) labor is relatively scarce? If so, it should be associated with somewhat higher wage growth for a given or growing labor stock and a reduction in unemployment. Thus, the hypothesis of this research is: *When the labor market becomes tighter, we expect higher entry wages for young individuals.* In detail, we raise three questions. First, can we identify higher individual wage growth for apprenticeship leavers, depending on the relative shortage in particular occupations? Second, is this effect more pronounced for specific occupations such as STEM? Third, does the impact differ depending on the individual position within the entire wage distribution? For this purpose, we construct a labor market tightness measure following the wage-curve literature (e.g., Blanchflower and Oswald 2008) and estimate an individual's 10-year wage-growth equation in a setting of Mincer (1974) for apprenticeship leavers. We use an extensive German administrative database of young individuals for 1995–2014 provided by the German Institute for Employment Research. Because we are explicitly interested in the effect of the demographic transition, we need to focus on individuals directly affected by such population change. Considering, for instance, unemployed individuals after mass layoffs does neither take the demographic transition into account nor does it reflect the tightness-problem of young individuals. We, therefore, stick to the group of young individuals entering the labor market.

This paper is structured as follows. A theoretical motivation and the hypotheses are given in Sect. 2. Section 3 introduces the data and variable constructions. The econometric methods used in this study are very briefly presented in Sect. 4. The results are presented and discussed in Sect. 5, and finally, Sect. 6 summarizes and concludes the paper.

2 Theoretical considerations and hypotheses

Believing that occupational labor market tightness is driven by changes in product demand, technological change, or demographic change, we set up arguments from the wage-curve literature in basic macroeconomic settings (see Blanchflower and Oswald 2008).

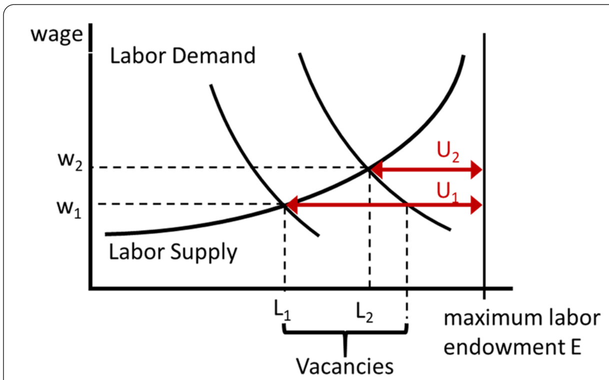

Fig. 1 Wages in a labor market demand and supply equilibrium

Consider the labor market in a specific occupation and region as outlined in Fig. 1. It visualizes the labor demand D and labor supply S curves. There is a maximum endowment of labor in a specific occupation and region, either employed E or unemployed U. In equilibrium, E_1 people are employed and U_1 people are unemployed, leading to wage w_1. We explicitly consider occupations instead of the entire regional employment-unemployment situation because of its heterogeneity. There is no reason to assume that the slopes of demand and supply side, potential shifts in both curves but also the maximum labor endowment, the effect of, e.g., occupational trends, the impact of technological change, etc., are identical over occupations. There might be (regional) trends in demand for specific fields, with demands rising for ones but falling in others. For a given stock of employees in both professions, the unemployment rate may fall in one occupation but increase in the other. Using aggregate regional data on unemployment would not be able to disentangle the different directions caused by aggregation. Depending on the relative size of trends in both occupations, we may observe a reduction or an increase in unemployment and, consequently, differences in wage reactions. Finally, in the public debate, the labor shortage is often claimed to exist for specific occupations, making this level of analysis relevant. The occupation-specific regional setting captures such heterogeneity and allows a much better understanding of the functioning of regional, occupation-specific labor markets.

In the last two decades, firms have often claimed that they had to reject business opportunities because of the lack of labor. We interpret this as a rising product demand associated with increased labor demand. If labor demand in an occupation increases, the demand curve shifts outwards. Wages are expected to rise from w_1 to w_2. In the short run, employers usually pay w_1; therefore, an excess demand creates vacancies. At the same time,

U_1 unemployed workers are available but are unwilling to take a job for a small wage w_1.

Within a labor market matching framework, the number of unemployed people and vacancies are usually utilized to construct a labor market tightness measure. However, only very scarce data on open vacancies is available at the regional-occupational level. Similar to Moscarini and Postel-Vinay (2017), we use the employment level E_1 instead. Hence, our tightness measure relates to the unemployed-to-employed in a region and occupation. In the medium and long runs, wages will adjust, and the "$E_2 - w_2 - U_2$"– combination will describe the new equilibrium. As can be seen, the labor market becomes tighter because the number of employees rises, whereas the number of unemployed people shrinks. Moreover, the question arises whether there are different effects on wage growth because of labor market tightness between income groups in the same occupational field (e.g., IT, sales, social occupations). Acemoglu and Restrepo (2018) present mechanisms through which human work is substituted or is complementary to capital-intensive production, depending on production costs and whether it leads to higher unemployment and stagnant wages for specific groups in the wage distribution. It is assumed that low-paid employees in a particular occupation who are hypothetically low-skilled can be easier to replace with capital (machines) than well-paid (high-skilled) employees in the same professional field. Thus, the bargaining power of the already well-paid employees is more vital than that of the low-paid people, which leads to rising wage inequality between these two income groups. Suppose fewer free/unemployed workers are available but not yet substitutable. In that case, lower-skilled workers might become more desirable again, and their wages might grow more—compared to better-skilled workers within that occupation—to prevent them from moving into other occupations that offer better conditions. Thus, the effect of tightness might differ between income groups within the same occupation.

Based on these theoretical reflections, we attempt to investigate the already-mentioned hypotheses:

H1: *Labor market tightness positively affects individuals' wages*;

H1: *The effects of labor market tightness differ between occupations*;

H1: *The effects of labor market tightness differ depending on the relative position of individual wages in the entire wage distribution.*

3 Data and sample selection

We use the individual data of the Integrated Employment Biographies (IEB) provided by the Institute for Employment Research (IAB) for 1995–2014. This administrative data includes all working individuals subject to social security contributions and those who are unemployed. It allows the construction of the entire (un-)employment history of about 90% of the entire German labor market. From the whole universe of employment entries of all individuals (so-called *spells*) contained in the IEB, we aggregate data at the level of the firm and all higher levels of hierarchy, such as the industry, occupation, and region, and derive a linked employer-employee dataset. From this data source we derive various measures at the individual level, such as information on the unemployment history, migration, and job-changing behavior, and also at the firm and region levels and within occupations. A comprehensive description of all variables, which we include in all estimations, is provided in Appendix 1: Table 4.

Card and Lemieux (2001) argue that different cohorts are imperfect substitutes because the older workers have already attained experience and thus are potentially more productive. On the other hand, the newest cohorts possess the latest knowledge but have almost no experience. They enter the labor market under similar conditions and have more or less equal productivity levels. Thus, individuals from younger generations are easy to compare because of their simple and short employment biographies that might affect their individual wages. This is why we consider a 10% sample of employees under 30 years of age who have just completed vocational training. Our homogeneous sample consists of about 350,000 individuals.

Although wage data is highly reliable, we drop a few outlier cases with exceptionally low or high wages. As we focus only on young individuals, their wages are below a specific social-security contribution threshold, and we do not have to impute truncated wages. Occupations are grouped into 54 distinct occupational fields based on the German classification of occupations (KldB-88) at the 3-digit level. This classification considers within-occupation mobility, such that lower mobility rates between occupations are included.

Because we consider the wage growth of the individuals between the first job and 10 years later, we only consider individuals who are employed at both moments in time. We are aware of the potential selection bias, especially for not controlling for being unemployed or excluded[1]

[1] There are several reasons, why individuals may drop out from the data basis: being self-employed, civil servant, on maternity leave, outmigration to another country. We explore the selection issue more carefully in the robustness section.

10 years after the first job. Our final sample consists of $N = 316,711$ individuals.

Our focus variables describe the economic conditions of the occupation-specific regional labor market. We assume more substantial individual wage growth in occupations and regions where relatively few potential workers are available to start a new job relative to the number of employees. Therefore, the focus variable *ue* is computed as the number of unemployed people divided by the number of employees within an occupational field and labor market region where an observed individual started to work after finishing vocational training. In addition, we compute the change, Δue, between the year of starting the first job and 10 years later.

To account for alternative explanations, we control for the occupation-specific regional labor market size to measure Marshallian Externalities (Combes et al., 2004). We also consider direct competition through the proportion of academic employees (holding a university degree) within an occupation and region to consider potential substitutionary or complementary relationships between apprenticeship and university-degree holders. In addition, the academics' proportion controls for skill-biased-technological-change (SBTC). If there is, for example, an increase in labor demand caused by SBTC, then the number of university-degree holders increases. For a given labor market size, fewer workers have passed vocational training. The second group is our focus group, making the labor market tighter from a vocational training perspective as their qualification level becomes scarcer. Furthermore, we include firm-specific characteristics at the entry time point and changes in characteristics such as firm size, firm age, the share of human capital, proportion of women, and foreigners within firms, while considering different productivity levels. Firm controls are essential determinants for wage differentials, as has been shown by Brunow and Jost (2022) and Dostie et al. (2020). A list of variables under consideration is provided in Appendix 1: Table 4.

4 Methodical approach

Inspired by Mincer's (1974) wage equation, we estimate monthly log-wages 10 years after the apprenticeship training through initial characteristics at the first job and the changes in these attributes.

We now define and formalize the setup with a model. Having the initial wage w_i^0 of an individual i entering the labor market in year t in a specific occupation o, firm f, region r as well as industry s, we intend to model the wage 10 years after, w_i^{10}, as

$$\begin{aligned}\log w_i^{10} = &\alpha_0 + \alpha_1 \log w_i^0 + x_i^0 \beta_1 \\ &+ \Delta x_i \gamma_1 + Z_i \beta_2 + \Delta Z_i \gamma_2 \\ &+ ue_{r_i,o_i} \beta_3 + \Delta ue_{r_i,o_i} \gamma_3 + v_{f_i,t_i} \beta_4 \\ &+ \Delta v_{f_i,t_i} \gamma_4 + \mu_{o_i} + \mu_{r_i} + \mu_{s_i} + \mu_{t_i} + \varepsilon_i. \end{aligned} \tag{1}$$

Computing a wage growth measure as the dependent variable requires controlling for not only initial conditions—e.g., in the specification of the growth model from Solow (1956)—but also for events occurring during the 10-year transition period. According to the wage equation from Mincer (1974), individual characteristics x_i^0 at the entry time point as well as their changes $\Delta x_i = x_i^{10} - x_i^0$ within the 10 years explain the individual wage. Furthermore, the vectors Z and ΔZ include control variables capturing characteristics at all levels of the hierarchy, such as the firm or the occupation-specific regional labor market situation, and their respective 10-year changes. The focus variables ue_{r_i,o_i} and $\Delta ue_{r_i,o_i}$ represent the regional labor market tightness (*ue*-ratio), which indicates the number of unemployed compared to employed individuals within the same occupation and region in the entry year and its 10-year time change, respectively, for individual i. This addresses Hypothesis 1. An interaction effect of both variables with occupational indicators takes occupation-specific heterogeneity in slopes into account to address Hypothesis 2. In addition, we control through dummy variables for unobserved heterogeneity within an individual's occupation μ_{o_i}, region μ_{r_i}, industry μ_{s_i}, and entry year μ_{t_i}.

The 10-year period is arbitrarily chosen. For instance, collective agreements or payment schemes may be foreseen and thus, bias our results. However, such agreements usually exist within occupations and industries, which we control and therefore take care of between occupations and between industry variations. Additionally, changes in individual wages are less likely to occur instantaneously because of changes at the aggregated labor market level. As a result, shorter periods are likely to be less valid for analysis.

We assume no correlation between ε_i and $\log w_i^{10}$ and can therefore consistently estimate the parameter vector $\eta = (\alpha_0, \alpha_1, \beta_1, \gamma_1, \beta_2, \gamma_2, \beta_3, \gamma_3, \beta_4, \gamma_4, \ldots)^T$ by ordinary least-squares (OLS). Because the sample is relatively homogeneous with respect to age and labor market experience, and the focus variable at the occupational-regional level is likely to be uncorrelated with individual wages, we treat the focus variable as exogenous. An estimation of the model including the occupation, region, time, and industry effects together

Table 1 The effect of the occupation-specific labor market tightness on individual wages

	Baseline		Full model	
	OLS	Ada LASSO	OLS	Ada LASSO
	(1)	(2)	(3)	(4)
ue-ratio effect	− 0.256 (0.022)	− 0.276 (0.015)	− 0.101 (0.032)	− 0.250 (0.015)
Δ *ue*-ratio effect	− 0.215 (0.015)	− 0.219 (0.014)	− 0.065 (0.048)	− 0.120 (0.012)
Control variables	Yes	Yes	Yes	Yes
Occupation, region, industry, and time FE	Yes	Yes	Yes	Yes
Occupation # (*ue* and Δ*ue*) interaction effects	No	No	Yes	Yes
N	316,711	316,711	316,711	316,711
Adj. R^2	0.589	0.589	0.593	0.591
No. of variables	324	132	520	186
AIC	392,069.1	392,327.6	390,087.0	390,346.4
BIC	395,407.5	393,746.1	395,643.9	392,340.9

Dependent variable is the individual wage 10 years after labor market entry

OLS estimation: baseline model without occupation-specific interaction effects (Columns 1 and 2); full interaction model using occupation-specific interaction effects for *ue* and Δ*ue* (Columns 3 and 4); adaptive LASSO estimates are reported in Columns 2 and 4, respectively

Most of the control variables are included after LASSO. Heteroscedasticity robust clustered standard errors at the level of region#occupation of the coefficients are in parentheses

with the focus variable(s)—excluding all other information—would indicate little correlation and thus endogeneity bias when the effects do not differ significantly. This test was performed and showed the desired pattern. Another argument supporting our assumption of no correlation is the fact that we observe initial entry wages w_i^0 that are by definition unaffected by autocorrelation. To address the issue of different aggregation levels (e.g., occupational, regional, industrial-specific, etc.), standard errors are clustered at the regional-occupational level.

Because of the sample size, variables tend to be significant. However, their explanatory power is potentially low, and the influence is negligible from an economic point of view, i.e., there are not just a few Euro-cent differences in gross wages. Additionally, the variable set is large, containing 628 variables with 291 binary variables, leading to typical collinearity problems during the OLS estimation. Therefore, some groups of variables tend to be statistically significant in separation, leading to a potentially nonexistent economic interpretation, but their joint power vanishes due to high correlation. For these reasons, we use the adaptive Least Absolute Shrinkage and Selection Operator (adaptive LASSO) from Zou (2006) to shrink the dimension, avoid collinearity in the structural parameters, and identify only those variables that offer explanatory power. A more detailed description of LASSO is provided in Appendix 2. The LASSO procedure then offers the advantage of reporting variables that turn out to be not just statistically but also economically significant. By economic significance, we mean a situation in which a change in wages is more than a few Euro-cents caused by a change in some explanatory variable.

In Hypothesis 3, we suspect that the *ue*-ratio affects individual wages and thus the whole wage distribution $F_{w_{t+10}}$ differently. For example, a changing labor market situation might influence the demand for low-paid jobs more than better-paid jobs and thus their wage growth within an occupation. Unconditional quantile regression is estimated to take this heterogeneity in the effects into account (see Firpo et al. 2009). Details on the estimation procedure can be found in Appendix 2.

5 Results

5.1 Does labor market tightness affect individual wages?

The estimated coefficients for ue and Δue are presented in Table 1. Column (1) shows the OLS estimates not controlling for occupation-specific heterogeneity in slopes (Hypothesis 1). Column 2 depicts consistent OLS estimates of the model with the most important variables selected by the adaptive LASSO approach. As can be seen, the set of explanatory variables is reduced by almost 200 parameters. To address Hypothesis 2, Column (3) shows the full model estimated with OLS, including 520 independent variables, and reports the selected model by using adaptive LASSO from Zou (2006) in Column (4). The latter model contains only 186 regressors. This again corresponds to a dimension reduction of two-thirds without a loss in the model fit. Regarding construction, the reduced model has fewer collinearity problems, particularly when considering interaction variables. The estimated coefficients are more precise because regularization has removed variables leading to the drastic increase in the variance. In addition, the dimension reduction serves as a kind of robustness check for the coefficient's significance. It is worth noting that almost all control variables remain in the models after LASSO, indicating the importance and explanatory power of the economic theory-led characteristics. The reduction in estimates is because of the exclusion of mainly regional indicators. Thus, most regions have no specific effect. Another drop relates to interaction effects, as we will discuss soon. All models can explain approximately 60% of the variance. Because of the advantages of the LASSO, these results will be interpreted, and the less precise estimates of OLS are shown for comparison only.

The coefficients of the unemployment variables in all model variations are negative and significant, as expected. The elasticity of $ue = -0.276$ (Column 2) is the mean effect over all occupations, which can be interpreted as follows: 1% of the decrease in the unemployed to employed ratio within the region and occupation at the first job would lead to an increase in log-wages of about 0.276% 10 years later. Considering Δue, a 1% reduction in the ue-ratio in the first 10 years after entering the labor market leads to an additional wage premium of approximately 0.219%. Nijkamp and Poot (2005) report a wage-curve elasticity of − 0.07, which is larger than our findings. For the mean ue-ratio, an elasticity of − 0.033 results in our case.

The wage-curve literature estimates the elasticity using current wage and unemployment levels (or lagged by one year). However, we consider growth and current unemployment's effect on future wages. This is why we expect lower elasticities, as current shocks/events typically have fading effects. Estimates of the German wage curve report smaller values (e.g., Baltagi et al. 2012). Interestingly, the study of Baltagi et al. (2012) reports a long-run elasticity of − 0.039, which perfectly aligns with our findings. In addition, we consider young individuals who are typically more flexible (see Axelrad et al. 2018) and therefore less likely to be affected by unemployment. Firms decide in favor of younger people; thus, they have greater bargaining power, especially if there is a labor shortage. Therefore, the effect of unemployment on young people's wages is less pronounced.

Our results are pretty robust through all our specifications and support hypothesis:

H1: Labor market tightness positively affects individuals' wages.

5.2 Differences between occupations

Since the first hypothesis is valid, we interact ue and Δue with occupation dummies to test the second hypothesis. The results are presented in Table 1, Columns 3 and 4, where the reported estimates relate to the reference group ("agricultural, husbandry,..."). To overcome this disadvantage, the occupation-specific effects of the OLS (bold black line) and adaptive LASSO (gray bars, separated by STEM and non-STEM) estimation are visualized in Fig. 2. Insignificant estimates in this and all proceeding figures take the value of the reference group for both LASSO and OLS. In this and the following plots, non-significant coefficients of the interaction variables can be interpreted as profession-specific overall effects of the same size as the reference group. Interestingly, the point elasticities for most occupations are negative and insignificant. Contradictory to theory, they become positive only for healthcare-related rather than manual occupations.

The LASSO algorithm does not have a specific feature for dummy variables that belong together. When LASSO decides to drop one variable, then it is excluded. Concerning content, it means that this dummy variable is not significantly different from the omitted reference group. Especially when there is a meaningful order in dummy variables (e.g., age groups) and some dummy drops out, the results of LASSO may become arbitrary. We carefully went through the set of included variables of LASSO and compared it with the complete list of OLS estimates. A disadvantageous selection did not appear. Roughly speaking: when OLS estimates were insignificant, they were usually selected by LASSO to be excluded (four variables). The interpretation does not change. Additionally, three significant control variables were excluded, but with content, they did not explain much when going into a detailed analysis. This is a very important result, as it shows the validity of the model and the specification and the impact in reflecting the arguments to include the factors. Thus, the LASSO algorithm mainly excludes industry, region, and occupation dummy variables. Concerning content, there is the "average" occupation, with "standard" effects and some occupations where we significantly differ from the "norm". These occupations are highlighted in the figures.

Focusing on the most negative effects shows a clear tendency towards relatively higher wage growth in STEM-oriented occupations, especially in mechatronics, energy electronics, and electricity but also in vehicle and aircraft construction. These occupations cover relevant tasks in innovative processes and drivers of technological

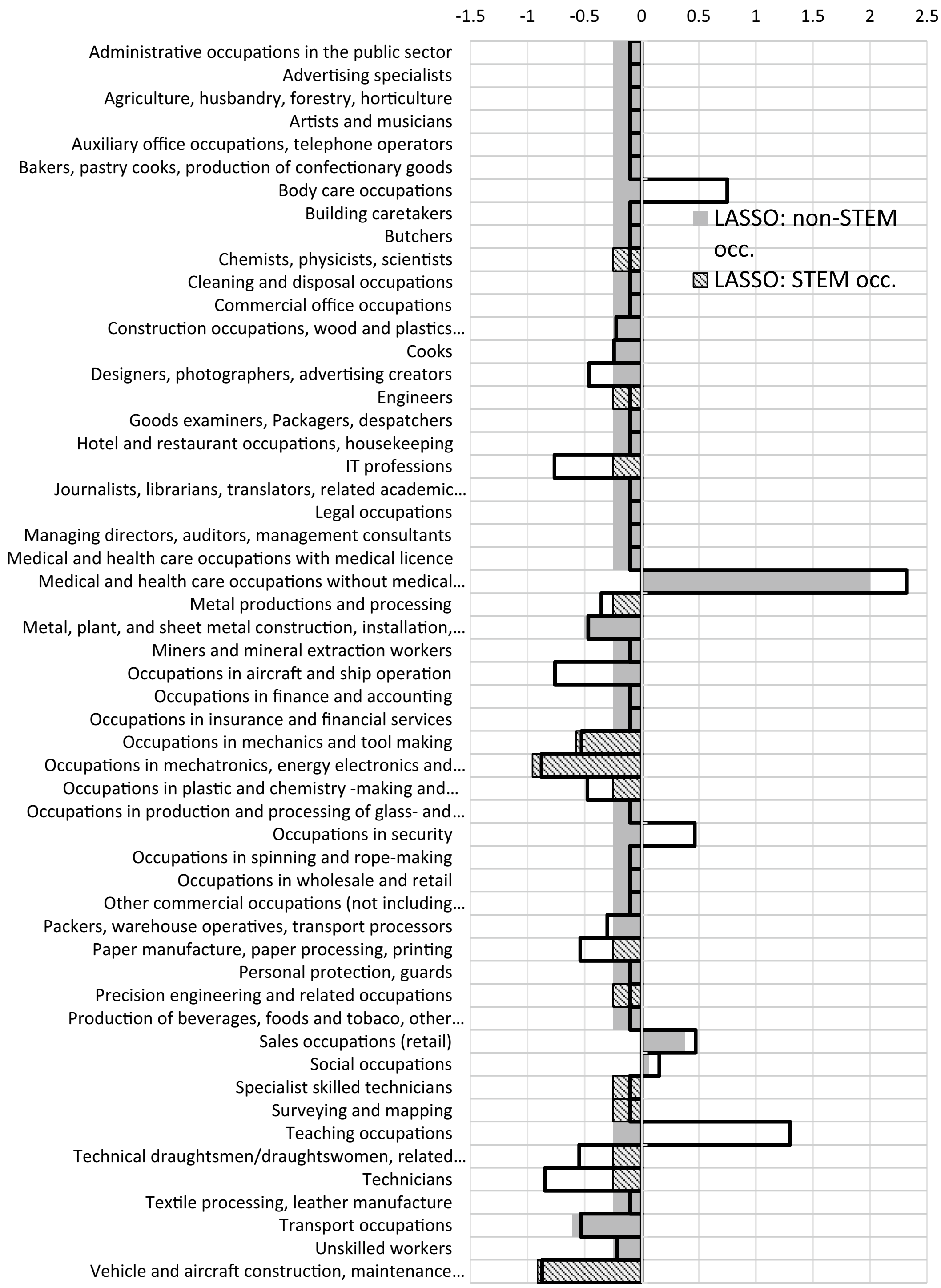

Fig. 2 Occupation-specific estimates of the *ue*-ratio on individual wages (sorted alphabetically)

change. For instance, mechatronics is a good example of the shift from "pure" mechanics to the combination of mechanics, electronics, and AI techniques. However, for the other STEM occupations, the effect is comparable to the other non-STEM occupations. Still, wages react most sensitively, especially in professions where a shortage is claimed. Therefore, employers respond to the shortage and indeed pay relatively higher wages when the labor market is tighter. In most cases, we confirm the negative effect of labor market tightness. "*Health care occupations without a license to practice medicine*" (e.g., nurses, geriatric nurses, etc.) are exceptions. The *ue* interaction coefficient is positive, significant, and very robust. The reasons for the unusual positive effect could be political interventions and the general demand for more health care due to the expected "Demographic change" and specific regulations regarding the health care system funding so that the suggested market mechanisms are not at work. In Germany, the health care system is organized in such a way as to prevent a "*two-classes-health-care-sector*": one class for those who cannot afford to pay for health care and the other class for those who can. For this reason, responsible central organizations, from health insurance companies to medical associations, negotiate fixed prices for medical treatments at the federal level. Thus, prices are aligned all over Germany with few state-level exceptions. Consequently, each medical treatment achieves the negotiated price; from this income, all costs must be covered. These costs cover wages but also other expenditures such as those for rent. Thus, because of the fixed prices, wages cannot react the way they do in a free market economy. Another regulation affects the number of doctors in a region. This is also regulated and mainly depends on the number of inhabitants in a region. However, existing medical associations decide whether there are sufficient "customers" to allow for another doctor to operate in that particular region. Thus, the number of "firms" is also regulated.

Figure 3 visualizes the effect on wages of the 10-year change in the tightness measure $\Delta ue_{r_i,o_i}$. The picture clearly shows the market reaction in health-related occupations, where wage growth is higher compared to other occupations. These occupations have been claimed to be the "missing" ones in Germany, and thus the market reacts. The Δue effect is still existing for some STEM occupations. In general, we can see that wages rose above average in all occupations when the labor market became relatively tighter because a significant effect is estimated for all occupations. Thus, our results suggest that employers react to labor market tightness and pay somewhat higher wages. The results are in favor of Hypothesis 2.

Worth mentioning is that not necessarily every variable chosen by the LASSO is significant, although some relationship between the *t*-test and shrinkage estimation in the simplest form can be made. The effect does not significantly vary relative to the reference group for either the *ue*-ratio or its 10-year change. This is not associated with the losses in information and interpretation but with a gain in efficiency and preciseness. The LASSO results identify occupations where a deviation from the general trend occurs and are therefore superior to the reported OLS.

5.3 Cost of tightness

Employers react to labor market tightness and pay higher wages when a shortage occurs. What does it mean in Euros? We take the occupation-specific median gross-monthly income and the distribution of the occupation-specific unemployment rate (centered at its occupation-specific average). We can compute the hypothetical wage increase for each occupation as shown in Fig. 4 based on model (4) of Table 1. For instance, consider an individual employed within the "*Occupations in mechatronics, energy electronics and electrical...*" in a region where the occupation-specific unemployment was relatively high while the individual was entering the labor market, i.e., at the 95th percentile of the entire distribution within this occupation. This individual will earn about €1,500. A person entering a region with hardly any unemployment in that occupation (5th percentile) enjoys a wage of €1,726, thus, about €226 more compared to a person working in a region with high unemployment (95th percentile). Again, in STEM occupations and some health-related occupations, wages react more sensitively the higher the shortage is when entering the labor market. The picture is inverted for the three occupations at the bottom of the table, which contradicts the theory—these are mainly healthcare-related occupations, as discussed in the previous section.

We now consider the change in labor market tightness and present the potential change in income due to relaxation or intensification at the regional-occupational level. The results are visualized in Fig. 5. For instance, in "*Occupations in mechatronics, energy electronics and electrical...*" with no change in the *ue*-ratio within the 10 years, there is zero additional wage growth (= zero line). However, when there is, for example, a sharp increase in unemployment (95%) such that the *ue*-ratio becomes more prominent, we observe a negative effect on wages of about €-72 relative to a region where no changes occurred. On the other side, when there is a relative reduction in unemployment (5%), such that the *ue*-ratio indicates a tighter situation (fewer unemployed individuals for each employed), employers pay about €45 higher wages relative to the zero line. This gives a wage range of €117, given the worst to best-performing regions within these occupations. Interestingly, here wage growth is not necessarily a STEM-related issue. However, it is worth noting that the *ue*-ratio does not vary that much within STEM occupations, and the regional range within each occupation is also smaller.

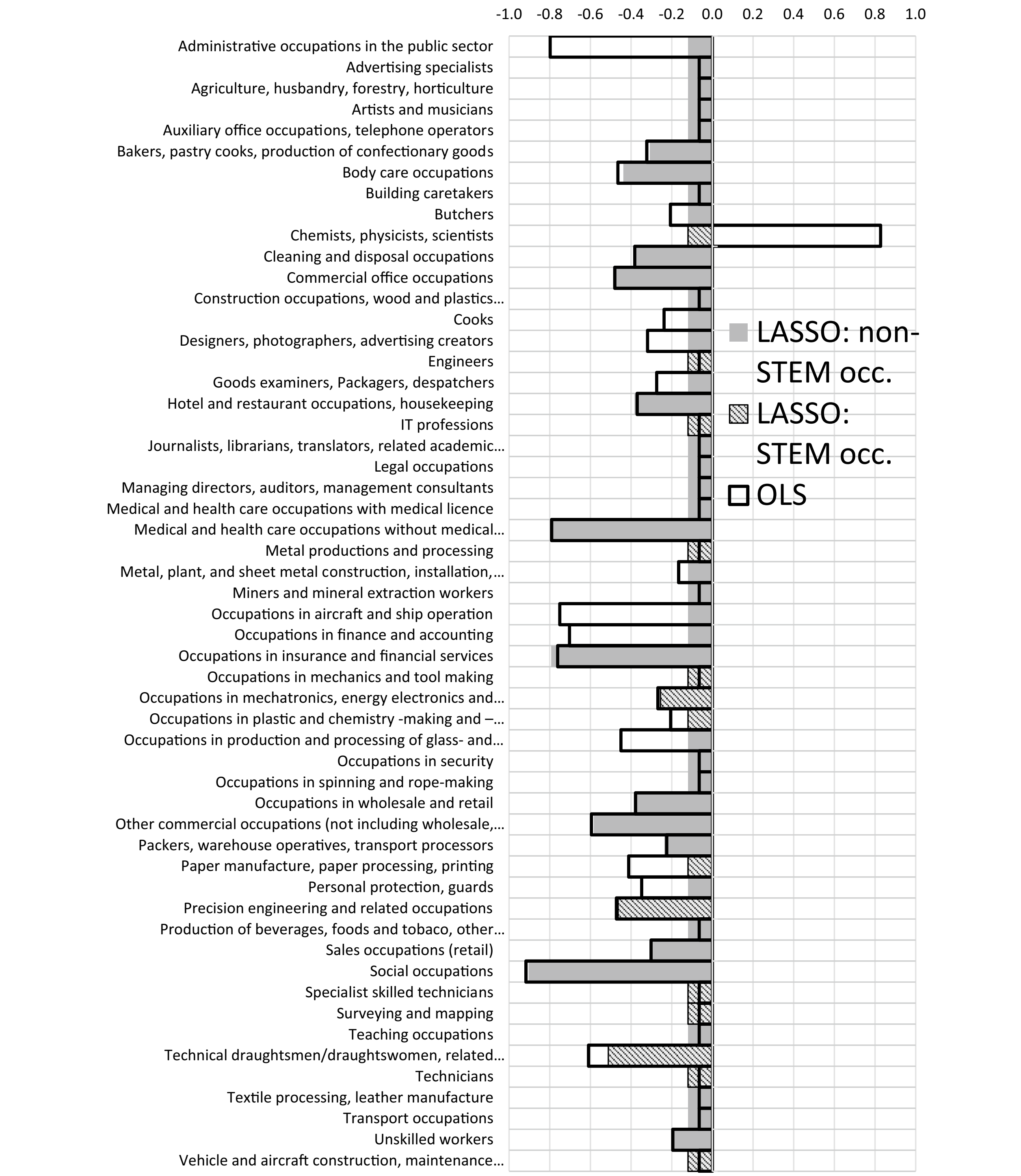

Fig. 3 Occupation-specific estimates of the growth of the *ue*-ratio on individual wages, sorted as in Fig. 2

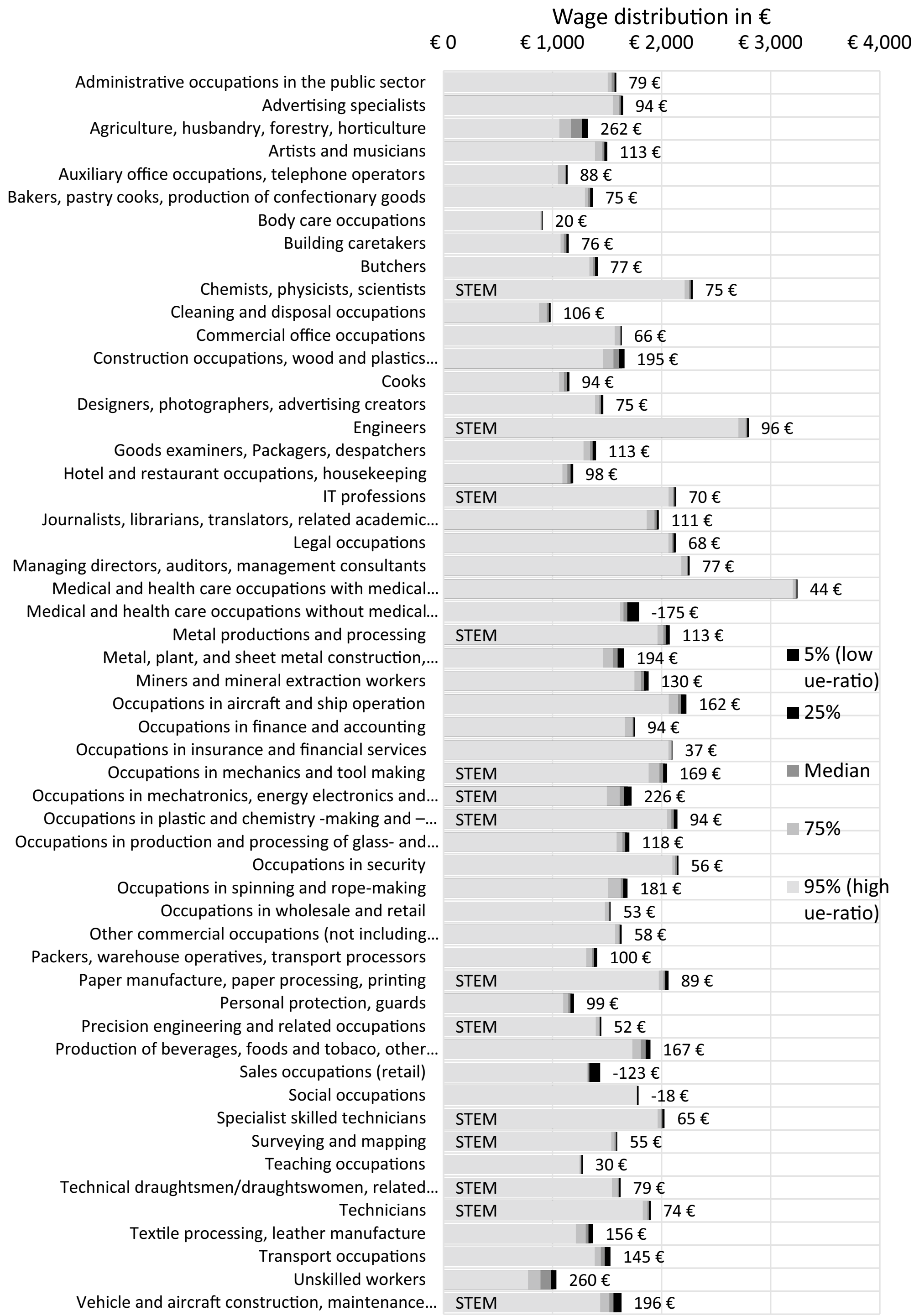

Fig. 4 Costs of Tightness (entry conditions of the *ue*-ratio)

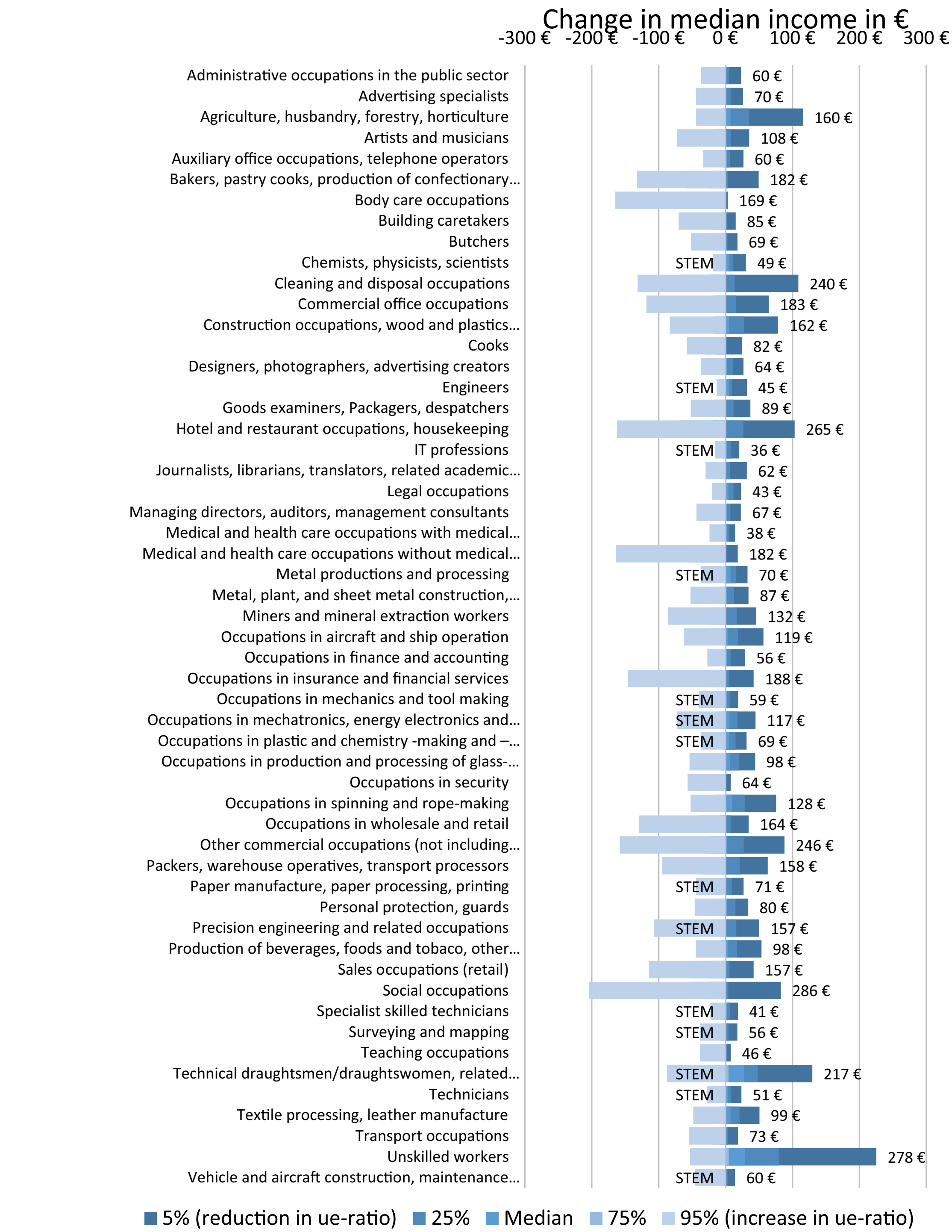

Fig. 5 Costs of tightness (10-year change in the ue-ratio)

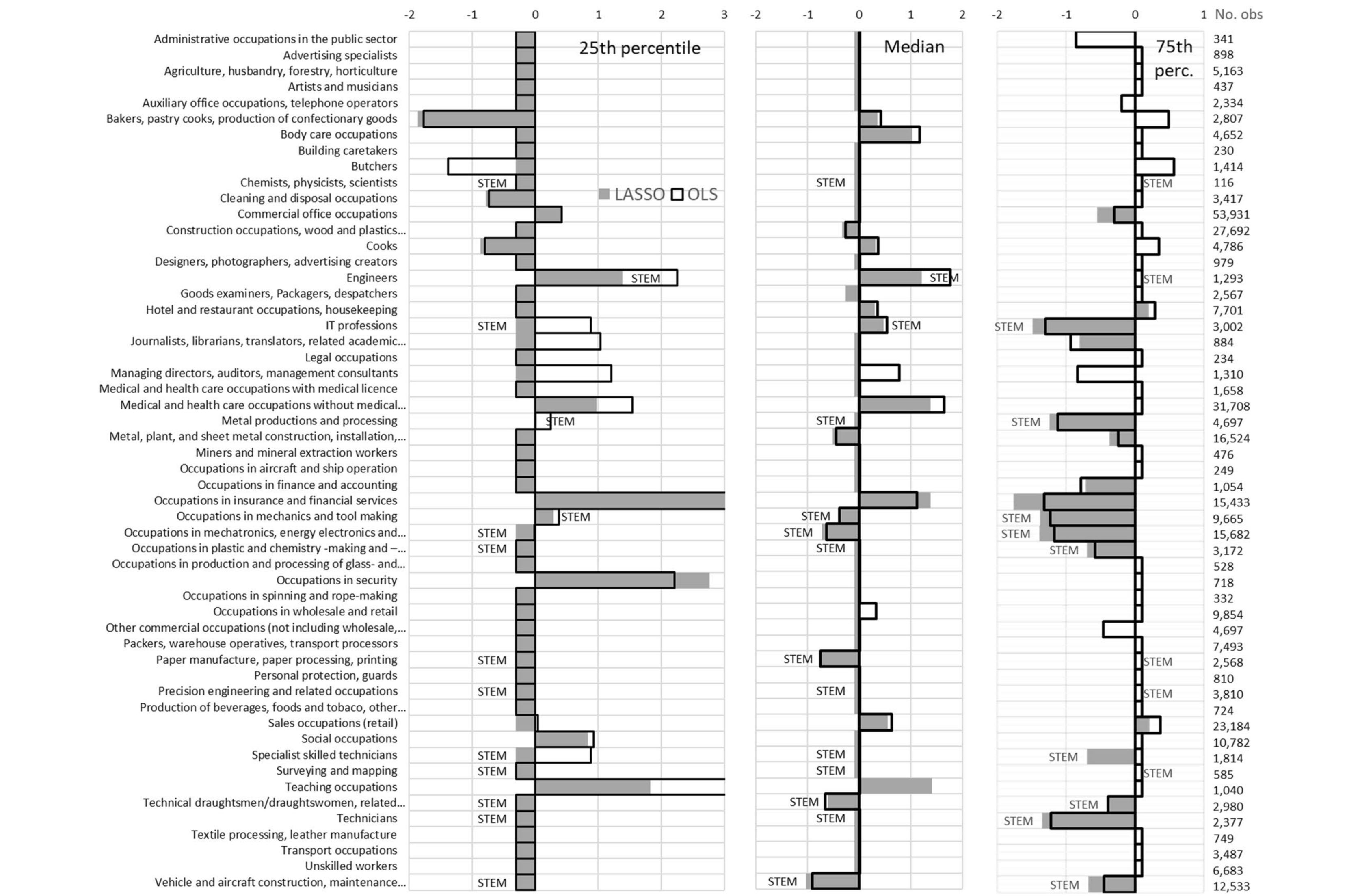

Fig. 6 Parameter heterogeneity of the *ue*-ratio (quantile regression). Notes: all effects include the main effect plus interaction effect, when the deviation is significant (95%); insignificant estimates show the parameter of the reference. Models include all control variables and occupation-specific effects for the *ue*-ratio (when selected by LASSO). Occupations sorted in alphabetic order

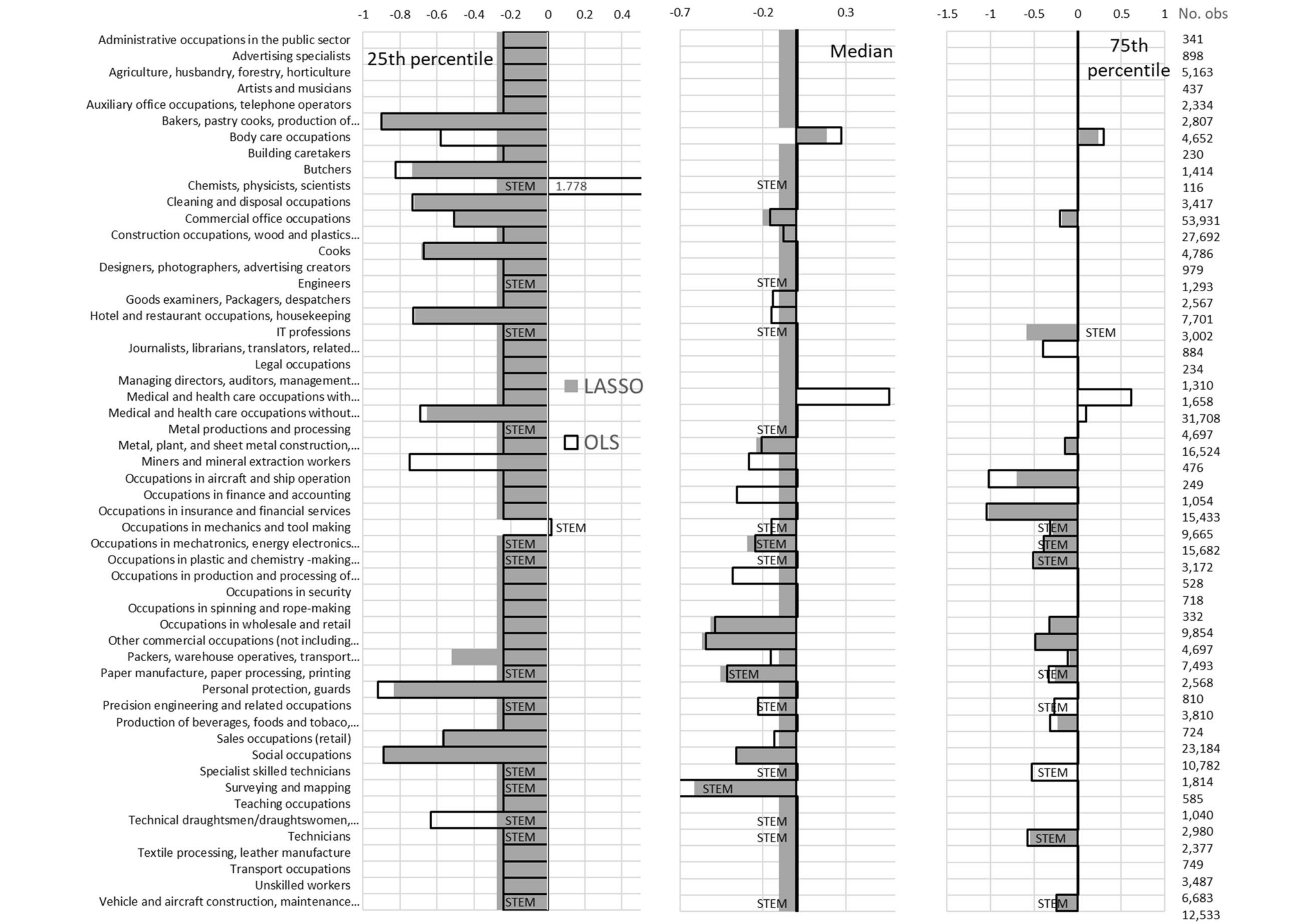

Fig. 7 Parameter heterogeneity for the ten-year change of the *ue*-ratio (Quantile regression). Notes: all effects include the main effect plus interaction effect, when the deviation is significant (95%), insignificant estimates show the parameter of the reference. Models include all control variables and occupation-specific effects for the ***Δ****ue*-ratio (when selected by LASSO). Occupations sorted in alphabetic order

Table 2 Results of the Quantile Regression according to Firpo et al. (2009)

	Evaluation of the wage distribution at the…					
	25th percentile		50th percentile		75th percentile	
	OLS	LASSO	OLS	LASSO	OLS	LASSO
ue ratio effect	− 0.301	− 0.311	− 0.021	− 0.099	0.101	− 0.011
	(0.083)	(0.032)	(0.034)	(0.018)	(0.022)	(0.010)
Δ*ue*-ratio effect	− 0.240	− 0.277	− 0.031	− 0.105	0.028	n.s
	(0.088)	(0.030)	(0.024)	(0.012)	(0.013)	
Interaction	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
N	320,006	320,006	320,006	320,006	320,006	320,006
Adj. R^2	0.446	0.445	0.452	0.451	0.386	0.385
AIC	695,379.6	695,767.8	357,553.0	357,772.8	429,077.4	429,262.4
BIC	700,941.9	697,721.6	363,115.3	360,569.9	434,639.6	431,974.1

Interaction model with occupational heterogeneity

n.s. not selected by LASSO

Heteroscedasticity robust clustered standard errors at the level of region#occupation of the coefficients are in parentheses

Our results on occupation-specific effects are pretty robust through all our specifications outlined and based on models estimated in the robustness Sect. 5.5. We can therefore support our second hypothesis:

H2: The effects of labor market tightness differ between occupations.

5.4 Differences between the quantiles of the wage distribution

We have presented evidence of tightness effects on average wages. However, each occupation has substantial wage variations. These variations result not just because of within-variations arising from aggregating 3-digit occupations into the occupational fields but also because of unobserved individual heterogeneity, among others. The ultimate question arises: Are there different effects on the log-wage quantiles through the changing labor market situations, as stated in Hypothesis 3? In layman's terms, do potentially more productive workers benefit more from tighter labor markets? Answering these questions, we estimate different quantile levels of the log wages across all occupations by using the unconditional quantile regression by Firpo et al. (2009). The independent variables for each quantile level $\tau = \{0.25, 0.5, 0.75\}$ are separately selected using adaptive LASSO for the model that considers occupational parameter heterogeneity as in Sect. 4 (see Figs. 6 and 7 for the *ue*-ratio and Δ*ue*-ratio, respectively, and Table 2). For comparison purposes, we also estimate the different quantile levels of the log-wages through all 628 regressors (Columns named OLS) and present them in the respective figures. Again, the model fit values remain the same, although almost two-thirds of the variables are dropped through the adaptive LASSO.

Considering the 25th percentile of the wage distribution, i.e., relatively lower wages, there is for most occupations a negative effect of the *ue*-ratio of about − 0.31. Interestingly, the effect is not stronger for the rather technical professions (i.e., STEM), indicating that lower-paid jobs in these occupations do not benefit from tightness above average. Wages react less strongly for the workers with median income. The effect is about − 0.099 and reflects the previously reported mean estimates. For the 75th percentile, the effect is less decisive. For most occupations, the estimate is zero; thus, irrespective of the degree of tightness, no additional wage premium on a tighter situation is observed (see Fig. 6). However, the strong negative impact for some occupations can still be observed, mainly in technical occupations such as STEM. Already mentioned exceptions, such as health-care occupations, are also found in the quantile distribution.

The change in the *ue*-ratio also shows heterogeneity between the quantiles of the wage distribution. Again, the effect is stronger for the lower-paid jobs and insignificant for the better-paid jobs. There are some occupation-specific effects to observe (see Fig. 7), especially social occupations and service occupations. One should be careful in explaining results from Figs. 6 and 7, as for some quantiles number of observations left for estimation are pretty small, as reported in the figures.

In the years under consideration, unemployment decreased steadily. Our results suggest that individuals, particularly at the lower end of the wage distribution, benefitted most from that positive development. The

Table 3 Estimates of focus variables in robustness checks

Model	Row	*ue*-ratio		Δ *ue*-ratio		No. Obs	Adj. R^2
		b	(s.e.)	b	(s.e.)		
Baseline	(1)	− 0.250	(0.015)	− 0.120	(0.012)	316,711	0.591
Men only	(2)	− 0.305	(0.017)	− 0.132	(0.014)	177,008	0.562
Women only	(3)	− 0.116	(0.035)	− 0.170	(0.046)	139,703	0.555
Educational degree	(4)	− 0.262	(0.016)	− 0.127	(0.012)	275,342	0.587
No University option	(5)	− 0.238	(0.017)	− 0.118	(0.012)	224,050	0.582
No regional migration	(6)	− 0.206	(0.019)	− 0.089	(0.012)	216,985	0.603
No migr., no STEM change	(7)	− 0.194	(0.020)	− 0.095	(0.014)	198,176	0.604
no STEM change	(8)	− 0.238	(0.016)	− 0.122	(0.014)	285,613	0.592
No reg. and no occ. Migr	(9)	− 0.189	(0.022)	− 0.148	(0.020)	124,752	0.594
Only one employer	(10)	− 0.008	(0.018)	− 0.055	(0.009)	316,711	0.636
Average firm wage	(11)	− 0.170	(0.028)	− 0.122	(0.018)	256,233	0.602
CHK effects	(12)	− 0.143	(0.025)	− 0.097	(0.029)	71,025	0.609

Each row represents one robustness check regression; cluster robust s.e. in parentheses

Including all control variables and occupation-specific interaction effects for the *ue*-ratio and Δ *ue*-ratio

wage reactions were stronger than those for the jobs at the other end of the wage distribution. Thus, lower-paid jobs became more attractive.

The derived results are in favor of Hypothesis 3:

H3: The effects of labor market tightness differ depending on the relative position of individual wages in the entire wage distribution.

5.5 Robustness checks

We perform substantial robustness checks on the tightness measure and on various subgroups of our sample, such as men or women, as well as those with or without educational degrees and with fewer mobility prospects.

We use the ratio of unemployed-to-employed people within occupations to measure tightness; the ratio of unemployed-to-number-of-vacancies is an alternative one (see Hershbein and Kahn 2018). By using the IAB-Job-Vacancy Survey, we also employ the number of vacancies. In this study, we are, unfortunately, limited to the years 2007 and 2008, in which occupation-specific data of some but not all occupations can be matched uniquely. The data limitation permits the construction of the 10 years of changes in vacancies. The results suggest no difference between the previously reported $(\Delta)ue$ coefficients, meaning that the *ue* variables capture the labor market tightness situation very well. The (log) vacancy coefficient is positive and significant, but with an effect of 0.009, it is tiny. It confirms that higher wages are expected if fewer unemployed people are available and the number of vacancies increases.

As another robustness check, we consider entry-cohort effects and estimate the model year by year to capture still uncontrolled heterogeneity among cohorts and years. The economic shock caused by the crisis in 2008/2009 and its policy and market reactions, for instance, may contradict the usual market mechanism in the short run, leading to high unemployment without any effect on wages in the long run. The Social Codebook reforms in 2004 may have changed the behavior of individuals during their life courses, and this, again, could impact the wages of entry cohorts differently. As shown in Appendix 1: Fig. 8, however, the effects of the *ue*-ratio are relatively stable over time and usually in the range of the joint estimate (the blue line).

Heterogeneity among subgroups might influence the effects, so we consider various subgroups. The results of the slope-augmented regression model (after LASSO) are presented row-by-row in Table 3 and only show the main effects of the *ue*-ratio and Δue.

Row 1 of Table 3 contains the baseline model for comparison. We estimate the model for men and women separately (rows 2 and 3). The *ue* effect is stronger for men than women due to the selectivity into gender-specific occupations. However, for women, there is a significantly higher estimate in Δue, indicating better labor market development.

The decision of an individual having additional training and/or changing a job and/or being regionally mobile can be driven by insufficient wages or expected positive career paths, which lead to a potential endogeneity problem. For instance, people are more likely to move when a job offer in another region is sufficient to cover all monetary and non-monetary costs. Thus, we observe higher wage growth. But this decision is made under the condition of income growth. For this reason, we only consider individuals with reduced regional and occupational mobility and/or fewer options for individual development within the 10 years (e.g., no option to study because of the educational level attained). We, therefore, consider more homogenous subsamples. Row 4 shows the results by limiting the original sample to individuals with a valid school-leaving certificate and completed vocational training. These individuals cannot enjoy higher wage growth than individuals without a degree because the latter still have the open option to obtain formal qualifications within the 10-year employment period. Individuals holding a secondary school-leaving qualification have no option to attend a university. They are, therefore, less flexible in the labor market as they have no "outside option" (see row 5). In both cases, the differences from the initial results are hardly visible, and the main findings hold.

Higher wage expectations can drive regional mobility; therefore, we exclude all regionally mobile individuals (row 6). Furthermore, we consider individuals who did not change from a typical STEM to a non-STEM occupation and vice versa (row 8) and joint regional and STEM mobility (row 7) to account for potential income growth perspectives and job mobility. The results hardly change if we make the sample even more homogeneous by excluding individuals who switched between occupations and regions (row 9). We restrict the sample to employees who do not change employers as another source of endogeneity driven by job-changing mobility during their life course (row 10). The main $(\Delta)ue$ coefficients become weaker in that case.

The results even hold if we augment the baseline model of the full sample with potential endogenous variables, i.e., average firm wages (and their changes) in row 11 and Card-Heining-Kline firm effects (row 12), which capture all unobserved firm heterogeneity. The main $(\Delta)ue$ coefficients become weaker, but the conclusion remains the same.

To account for occupation-specific productivity growth and technological change's potential effects on occupations' relevance, we added year-specific occupational indicators that absorb all-time trends within occupations. Again, the results provide the same conclusion.

Lastly, there is a selection bias issue, as we only include individuals, which are employed in both moments in time. Models, such as the Heckman approach, need at least one variable, which relates to the selection but not to the wage equation. Unfortunately, all of the variables, we observe, are relevant in the wage equation and thus, a Heckman procedure cannot be applied. For this reason, we run a multinomial logit model on the probability to be either included after 10 years or to be unemployed, or to be missing in the data (self-employed, maternity leave, civil servant, outmigration). The data set shows that about 10% are unemployed 10 years later and another 20% have missing information. The multinomial logit provides insights into the potential group specifics:

- The probability to be unemployed or being "dropped out" reduces from year to year, which can be explained by the general economic improvement in the years under consideration.
- There are some region, industry, and occupation-specific effects, which are, however, not too pronounced, i.e. the effects are rather small in magnitude.
- Females, mothers, and foreigners are more likely to be unemployed. Their propensity to drop out, however, is even higher.
- Individuals without a school degree and who did not successfully finish their apprenticeship, who spent more time in occupational training and were looking for the entry job relatively longer after occupational training compared to others, and those, who were of age less than 18 years when entering the first job, have a significantly higher propensity to be unemployed after 10 years. All these characteristics indicate, that these individuals have potentially their difficulties in the labor market and are therefore more affected by unemployment.
- In contrast, those individuals, who dropped out, are relatively better skilled, had a significantly shorter occupational training period, and have—compared to those, who are employed 10 years later—a slightly longer search time between the apprenticeship and the first job.
- With respect to our focus variables, we see that an increase in the *ue*-ratio and a larger labor market make it more likely to be unemployed and less likely to drop out. However, disentangling the effects by gender show, that an increase in the *ue*-ratio does not explain a higher propensity to be unemployed 10 years after entry for males but only for females. For both genders holds: if the labor market gets tighter (lower unemployment), the propensity to drop out increases.

The data set comprises almost half-half males and females. The propensity to become unemployed due to changes in the U–E-ratio is not significant for males and the effect is strong for females. It becomes more likely for females drop out after 10 years (probably because of maternity leave). The Multinomial Logit results indicate that drop-outs belong to the better skilled individuals. It is thus to expect that the bias in parameters is downward, as the better skilled individuals, which most likely would enjoy higher wage growth, drop out.

To sum up, the strength of *ue*-ratio depends on the occupations and subsamples, but the conclusion is always the same: Because of the reduction in unemployment in the past two decades, wages responded and rose in the past. Wages of STEM occupations usually react more sensitively to labor market tightness, indicating that technological progress associated with higher labor demand puts additional pressure on that field.

6 Concluding remarks

This paper investigates the ten-year wage development of young employees who had finished vocational training and started their first job and focuses on the effect of regional occupation-specific labor market tightness on individuals' wage growth. The underlying economic mechanism is that a workforce shortage affects wages positively if the labor demand remains constant or rises. Especially young individuals, who become relatively scarce in the future because of the population change, would benefit in tighter occupations. To verify this hypothesis, we measure labor market tightness by using the ratio of unemployed to employed people, *ue* -ratio, on occupational and regional levels. Then, we look for a tightness effect on the individual wages for 54 different occupational fields.

Therefore, the individual employment biographies (IEB) data are collected, including information on individuals before they enter the labor market during vocational training and the first 10 years after starting their first job. In addition, we match data with firm and labor market characteristics to the data on individuals and receive a linked employee-employer data set. Since we run a regression model with more than 628 independent variables, a variable selection by the adaptive LASSO approach from Zou (2006) is applied. The selected model is quite robust. The two focus variables, namely the *ue*-ratio at the first job and its change during the first 10 years after entering the labor market, show the expected negative signs.

The results indicate rising log wages when the labor supply becomes tighter. This effect varies between occupations and is stronger for some occupations, such as several engineer groups, technicians, IT, and commercial occupations. Thus, some typical STEM occupations have above-average benefits from labor market tightness. The results indicate a wage premium of up to 262€ within occupations in tight labor markets relative to a region with a "relaxed" labor market. Additionally, the change in the 10 years in labor market tightness matters and wage premia up to 270€ can be achieved. Thus, employers react to labor market tightness and pay substantially higher wages. Running an unconditional quantile regression, different effects in the lower (poor) and the upper (rich) quantiles become visible. Interestingly, the effect seems to reverse for employees in sales and especially healthcare occupations. The unusual positive effect could be political interventions and specific regulations regarding the health care system funding so that the suggested market mechanisms are not at work. These occupations need further investigation, which is not possible given our data.

To conclude, labor market tightness affects the individual wages of young individuals. Its strength depends on the occupations (and subsamples), but the conclusion is always the same: Because of the reduction in unemployment in the past two decades, wages reacted and rose in the past. Wages of STEM occupations usually respond more sensitively to labor market tightness, indicating that technological progress associated with higher labor demand puts additional pressure on that field. Individuals at the lower end of the wage distribution benefit relatively more from tightness.

Our results have important implications for the occupational choice of young individuals. Because of the competition for young individuals associated with higher wage growth, incentives for young individuals to choose such occupations are set. Especially STEM occupations favor this development, and thus, a structural change occurs. This development has costs: lower-paid occupations may face problems recruiting young labor. Suppose there is still a high demand for the products and services behind such occupations and no technical substitution possibility, at least in the mid- and long-run. In that case, immigration incentives should focus on all occupations to guarantee a sufficient stock of labor in all fields to compensate for the shrinkage of the entire workforce. For young individuals, our results suggest that a relative labor shortage in specific occupations and regions may lead to higher future wages and that occupational choice should take these implications into account besides individual interests and capabilities (see Table 4, Fig. 8).

Appendix 1

See Table 4 and Fig. 8.

Table 4 Description of the independent variables

Regressors	Range
Individual characteristics	
Age	Binary; classification: 15 to < 18 years, 18 to < 21 years, 21 to < 24 years, 24 to < 27 years and 27 to 30 years
Gender	Binary; male and female
Mother	Binary; whether a woman interrupts her job for child care within the 10 years of observation time
Foreigner	Binary; individual did not have a German passport
School-leaving qualification	Binary; no degree, lower secondary education (Hauptschule), secondary education (Mittlere Reife), higher education (Abitur)
Additional qualification	Binary; Master craftsman or university degree; interacted with school-leaving qualification
Duration of the vocational training	Binary; 3–4 years, 4 and more years
Period of time between the finishing of the occupational training and the start of the first job	Binary; 0 > duration > 3-month, 3-month ≤ duration < 1 year; duration ≥ 1 year
Log (wage)	Metric; log (average net wage per day) at the first job
Full- and part-time	Binary; individual works full- (part-) time at the first job and 10 years later changes from full- to part-time and vice versa
Individual mobility	
Switching from STEM to non-STEM and vice versa	Binary; whether an individual switched from a STEM to a non-STEM occupation and vice versa, classification from the German Federal Employment Agency (KldB 88–3 digit)
Upgrade/downgrade	Two discrete, ordinate variables; occupations are ranked with respect to their wages at the first job and 10 years after
Change the employer	Binary; changing the employer after the vocational training for the first job; changing the employer between the first job and 10 years later
Enforced firm change due to firm closure	Binary; whether an individual becomes unemployed due to the closure of the first employer's firm
Regional migration	Binary; whether individual moves to another labor market within the 10 years; additional interaction with STEM dummy
(log) Distance of regional migration	Metric; (log) distance of the labor market migration between the first job and 10 years later in km (interacted with regional migration dummy)
Migration from East to West Germany	Binary; migration from former East Germany to West Germany
Unemployment periods > 3 month	Discrete; the number of unemployment periods with > 3-month duration within the 10 years
Additional firms	Binary; an additional number of firms (from 1, 2, 3, 4, and 5 +) where individuals worked within the 10 years
Duration in employment	Metric, discrete; total number of days in employment (independent from employer) of individual i within the 10 years
Firm characteristics	
Firm age	Binary; whether the firm is up to < 5 years old
Log (firm size)	Metric; log (number of employees in the firm) at the first job
Share of women in the firm	Metric; at the first job
Share of foreigners in the firm	Metric; additional for robustness check; not in the baseline model
Share of human capital	Metric; the proportion of experts and specialists among all employees
Average firm wage	Metric; additional for robustness check
CHK effects	Card-Heining-Klein firm fixed effects (for robustness checks)
Occupation and regional characteristics	
Labor market size	Metric; log (number of employees in the occupational field and region of individual i) and interacted with an occupation dummy; both for time point t and the growth between t and $t + 10$ of individual i
Academic competitors	Metric; $\frac{numberofemployeesholdingauniversitydegree}{numberofallemployees}$ in the occupational field and labor market region of individual i and interacted with an occupation dummy; both for time point t and the growth between t and $t + 10$ of individual i
ue-ratio	Metric;$\frac{unemployedpeople}{employedpeople}$ in the occupation field and labor market region of individual i and interacted with an occupation dummy; both for time point t and the growth between t and $t + 10$
Labor market dummies	Binary variables; Labor market regions, according to Eckey et al. (2006)

Table 4 (continued)

Regressors	Range
Cluster heterogeneity and fixed effects	
Industry—Dummies	Binary variables
Time—Dummies	Binary variables

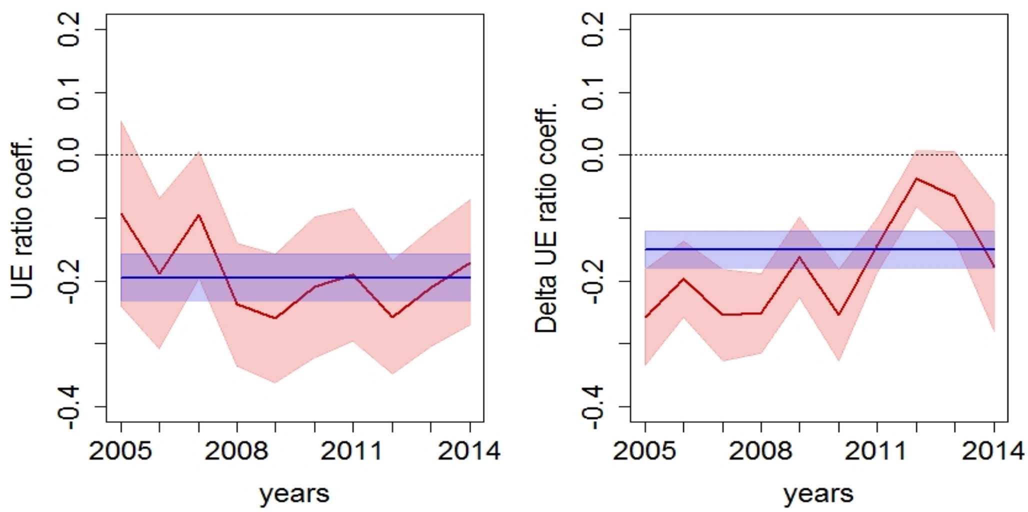

Fig. 8 Year-by-year (red) and the overall (blue) mean estimation of the *ue* and *Δue* coefficients using models selected by adaptive LASSO, including the 95%CI

Appendix 2

For simplification purposes, all variables in (1) are put together in z_t^T with a parameter vector η:

$$\log w_i^{10} = z_i^T \eta + \varepsilon_i \tag{2}$$

The idea is to penalize the parameters of RHS-variables (η), which hardly influence the LHS-variable $\log w_i^{10}$, and push them down to zero. The aim is to obtain a simpler model that does not include rather irrelevant variables. It should help to improve the forecast accuracy and the interpretability of the models. The default LASSO approach by Tibshirani (1996) results in the LASSO coefficient vector

$$\widehat{\eta}^* = \operatorname{argmin}_{\eta^*} \left\| \log w_i^{10} - \sum_{p=1}^{P} z_p \eta_p \right\|^2 + \lambda \sum_{p=1}^{P} |\eta_p|. \tag{3}$$

Thereby $\lambda \geq 0$ is the penalty coefficient, which is found by cross-validation, and p denotes a specific explanatory variable in the model. However, the LASSO variable selection can be inconsistent under certain conditions (see Meinshausen and Bühlmann 2006; Fan and Li 2001). For that reason, Zou (2006) augmented the penalty term through different weights for different coefficients. This yields

$$\widehat{\eta}^{*AD} = \operatorname{argmin}_{\eta^{*AD}} \left\| \log w_i^{10} - \sum_{p=1}^{P} z_p \eta_p \right\|^2 + \lambda \sum_{p=1}^{P} \frac{1}{|\eta_p|^{\gamma^{*AD}}} |\eta_p|, \tag{4}$$

with $\gamma^{*AD} > 0$ and $\lambda \geq 0$. The data-driven weights should ensure the so-called oracle properties, see Zou (2006).

Since the resulting LASSO coefficient vector $\widehat{\eta}^{*AD}$ is biased, we estimate the reduced model after the variable selection by OLS and compare the estimated coefficients of the main variables; this is in line with Chernozhukov et al. (2015).

For the quantile regression approach, we estimate the distribution $F_{w^{10}}$ of w^{10} through the unconditional quantile regression method by Firpo et al. (2009). The approach is based on the concept of sample quantile

q_τ added to the influence function $IF(w^{10}; q_\tau, F_{w^{10}})$ of this quantile of the wage growth for a specific level$\tau = \{0.10, 0.25, 0.50, 0.75, 0.90\}$:

$$\widetilde{w}_{\tau,10} := q_\tau + \frac{\tau - I\{w^{10} \leq q_\tau\}}{f_{w^{10}}(q_\tau)} = z_i^T \eta(\tau) + \widetilde{\varepsilon}_i, \quad (5)$$

where $\widetilde{w}_{\tau,10}$ is the unconditional τ quantile, which is called the recentered influence function, and $I\{w^{10} \leq q_\tau\}$ is the indicator function. The density $f_{w^{10}}(q_\tau)$ depends on q_τ and is estimated in our study using a kernel density estimator with the Gaussian kernel. We regress a set of variables z_t against $\widetilde{w}_{\tau,10}$. The coefficient vector η depending on τ can be simply estimated via OLS. Compared with the more traditional conditional quantile regression method by Koenker (2005), the estimation takes place in one step. Since the model is estimated via OLS, no particular assumptions for the error term vector $\widetilde{\varepsilon}_i$ are needed. There are no convergence problems. In addition, the marginal effect of a change in the distribution of z_t on the marginal quantiles of w^{10} is directly readable through the coefficients η. For further information, see Firpo et al. (2009) and Borah and Basu (2013).

Acknowledgements
Not applicable.

Author contributions
Authors equally contributed to writing, data preparation, development of the models, and analysis. All the authors read and approved the final manuscript.

Funding
No funding is applicable.

Availability of data and materials
The data used for this paper is confidential by German data security law and cannot be published. However, within a Guest Research stay at the German Institute for Employment Research, it is possible to access the data sets used. Data access requests a standardized formal approval by the German Federal Ministry of Labor and Social Affairs.

Declarations

Competing interests
There is no conflict of interest.

Author details
[1]Professor of Economics, University of Applied Labour Studies (HdBA), Campus Schwerin, Schwerin, Germany. [2]Chair of Statistics and Econometrics, Esp. in Transportation, Faculty of Transportation "Friedrich List", Technische Universität Dresden, Dresden, Germany.

Received: 28 February 2022 Accepted: 21 September 2022
Published online: 11 October 2022

References

Acemoglu, D.: Technical change, inequality, and the labor market. J. Econ. Literature **40**(1), 7–72 (2002)
Acemoglu, D., Autor, D.: Skills, tasks and technologies: implications for employment and earnings. Handb. Labor Econ. **4**, 1043–1171 (2011)
Acemoglu, D., Restrepo, P.: The race between man and machine: implications of technology for growth, factor shares, and employment. Am. Econ. Rev. **108**(6), 1488–1542 (2018)
Autor, D.H., Dorn, D.: Inequality and specialization: the growth of low-skill service jobs in the United States. NBER Working Paper Series 15150. SSRN J. (2009). https://doi.org/10.2139/ssrn.1434624
Autor, D.H., Levy, F., Murnane, R.J.: The skill content of recent technological change: an empirical exploration. Q. J. Econ. **118**(4), 1279–1333 (2003)
Autor, D.H., Katz, L.F., Kearney, M.S.: The polarization of the US labor market. Am. Econ. Rev. **96**(2), 189–194 (2006)
Autor, D.H., Katz, L.F., Kearney, M.S.: Trends in US Wage inequality: revising the revisionists. Rev. Econ. Stat. **90**(2), 300–323 (2008)
Axelrad, H., Malul, M., Luski, I.: Unemployment among younger and older individuals: does conventional data about unemployment tell us the whole story? J. Labor Mark Res. **52**(1), 3 (2018)
Baltagi, B.H., Blien, U., Wolf, K.: A dynamic spatial panel data approach to the German wage curve. Econ. Model. **29**, 12–21 (2012)
Blanchflower, D.G., Oswald, A.J.: Wage curve. In: Vernengo, M., Caldentey, E.P., Rosser, B.J. (eds.) The new palgrave dictionary of economics. Palgrave Macmillan, London (2008)
Borah, B.J., Basu, A.: Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. Health Econ. **22**(9), 1052–1070 (2013)
Brunow, S., Jost, O.: Wages of skilled migrant and native employees in Germany: new light on an old issue. Int. Migr. Rev. **56**(2), 410–432 (2022)
Card, D., Lemieux, T.: Can falling supply explain the rising return to college for younger men? A cohort-based analysis. Q. J. Econ. **116**(2), 705–746 (2001)
Chernozhukov, V., Hansen, C., Spindler, M.: Valid post-selection and post-regularization inference: an elementary, general approach. Ann. Rev. Econ. **7**(1), 649–688 (2015)
Combes, P.-P., Magnac, T., Robin, J.-M.: The dynamics of local employment in France. J. Urban Econ. **56**(2), 217–243 (2004)
Desjonqueres, T., Machin, S., Van Reenen, J.: Another nail in the coffin? Or can the trade based explanation of changing skill structures be resurrected? Scand. J. Econ. **101**(4), 533–554 (1999)
Dostie, B., Li, J., Card, D., Parent, D.: Employer policies and the immigrant-native earnings gap, IZA discussion papers 13245, Institute of labor economics (IZA), (2020)
Dustmann, C., Fitzenberger, B., Schönberg, U., Spitz-Oener, A.: From sick man of europe to economic superstar: Germany's resurgent economy. J. Econ. Perspect. **28**(1), 167–188 (2014)
Eckey, H.-F., Kosfeld, R., Türck, M.: Abgrenzung Deutscher Arbeitsmarktregionen. Raumforsch. Raumordn. **64**(4), 299–309 (2006)
Fan, J., Li, R.: Variable selection via nonconcave penalized likelihood and its oracle properties. J. Am. Stat. Assoc. **96**(456), 1348–1360 (2001)
Federal Employment Agency Germany (Statistik/Arbeitsmarktberichtserstattung): *Fachkräfteengpassanalyse,* Nürnberg (2017)
Firpo, S., Fortin, N.M., Lemieux, T.: Unconditional quantile regressions. Econometrica **77**(3), 953–973 (2009)
Goldin, C.D., Katz, L.F.: The race between education and technology. Harvard University Press, Cambridge (2009)
Goos, M., Manning, A., Salomons, A.: Explaining job polarization: routine-biased technological change and offshoring. Am. Econ. Rev. **104**(8), 2509–2526 (2014)
Hershbein, B., Kahn, L.B.: Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. Am. Econ. Rev. **108**(7), 1737–1772 (2018)
Koenker, R.: Quantile regression. Cambridge University Press, New York (2005)
Meinshausen, N., Bühlmann, P.: Variable selection and high-dimensional graphs with the Lasso. Ann. Stat. **34**, 1436–1462 (2006)
Mincer, J.: Schooling, experience, and earnings. Columbia University Press, New York (1974)
Moscarini, G., Postel-Vinay, F.: The relative power of employment-to-employment reallocation and unemployment exits in predicting wage growth. Am. Econ. Rev. **107**(5), 364–368 (2017)
Nijkamp, P., Poot, J.: The last word on the wage curve? J. Econ. Surv. **19**, 421–450 (2005). https://doi.org/10.1111/j.0950-0804.2005.00254.x
Solow, R.M.: A contribution to the theory of economic growth. Q. J. Econ. **70**(1), 65–94 (1956)

Spitz-Oener, A.: Technical change, job tasks, and rising educational demands: looking outside the wage structure. J. Law Econ. **24**(2), 235–270 (2006)
Statistisches Bundesamt (Destatis).: *Bevölkerungsentwicklung bis 2060 Ergebnisse der 13. koordinierten Bevölkerungsvorausberechnung,* Wiesbaden: Statistisches Bundesamt Deutschland (2017)
Tibshirani, R.: Regression shrinkage and selection via the Lasso. J. r. Stat. Soc. Ser. B **58**(1), 267–288 (1996)
Zou, H.: The adaptive Lasso and its Oracle properties. J. Am. Stat. Assoc. **101**(476), 1418–1429 (2006)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.