Polarization, Sorting and the Urban Wage Premium – Insights on German Local Labor Markets*

Katja Gehr[†] Valérie von Gleichen[‡]

March 8, 2024

Abstract

We establish three stylized facts on the evolution of spatial wage inequality, skill sorting, and employment polarization in Germany between 1980 and 2010. First, spatial wage inequality between high and low skill workers has increased over time. Second, spatial skill sorting has also increased. Third, employment polarization is stronger in denser cities and affects especially low skill workers. Based on this evidence, we develop a model that embeds the task-based framework into a spatial equilibrium model with heterogeneous workers to highlight the role of occupational change for the increase in spatial wage inequality and skill sorting. The key feature of the model is that the absolute and comparative advantage of cities and workers drive wages, city sizes, and sorting of workers across tasks and cities in spatial equilibrium. Overall, we find that routine-biased technical change leads to an increase in spatial wage inequality and skill sorting.

Keywords: Spatial Wage Inequality, Agglomeration Economies, Spatial Skill Sorting, Employment Polarization

JEL Codes: J24, J31, R23

*We thank Michael Pflüger for very insightful comments and discussions. We also acknowledge helpful comments and suggestions from seminar participants in Munich.

[†]Faculty of Economics, University of Würzburg, Sanderring 2, 97070 Würzburg, Germany, E-mail: katja.gehr@uni-wuerzburg.de.

[‡]Faculty of Economics, University of Würzburg, Sanderring 2, 97070 Würzburg, Germany, E-mail: valerie.gleichen@uni-wuerzburg.de.

1 Introduction

Since the 1980s, wage inequality between high and low skill workers has increased in many countries around the world, see Autor et al. (2008) for the US and Dustmann et al. (2009) for Germany. However, the increase in wage inequality has not evolved uniformly within countries but has an urban bias. Baum-Snow and Pavan (2013) and Baum-Snow et al. (2018) document the development of a positive relationship between wage inequality and city size. The increase in spatial wage inequality has been accompanied by a differential evolution of the urban wage gradient for high and low skill workers and a substantial increase in the geographic sorting of workers by skill. In a recent work, Autor (2019) suggests that labor market polarization can potentially explain why agglomeration forces have diverged across skill groups driving the increase in spatial wage inequality.

In this paper, we shed light on the role of occupational change for the increase in spatial wage inequality and skill sorting. Therefore, we establish three stylized facts on the evolution of spatial wage inequality, skill sorting, and employment polarization in Germany between 1980 and 2010. First, we find that the gap between the average wages of high and low skill workers has become more positively related to city density since 1980, i.e., spatial wage inequality has increased over time. This increase in spatial wage inequality is accompanied by an increase in the urban wage gradient for high skill workers, whereas the gradient for low skill workers declined, indicating a rise in the skill bias of agglomeration economies. The second fact is that the skill ratio of high to low skill workers has diverged across cities with denser cities becoming more high skilled (i.e., increased spatial skill sorting). Finally, we show that there is a spatial dimension to employment polarization with polarization being more pronounced in denser cities and affecting especially low skill workers.

Motivated by these facts, we develop a spatial equilibrium model to highlight the role of occupational change for the increase in spatial wage inequality and skill sorting. Therefore, we integrate the task-based framework in Acemoglu and Autor (2011) with the spatial equilibrium model of Diamond and Gaubert (2022). Building on the literature on the task-based framework, we consider a continuum of tasks that are produced with heterogeneously skilled workers and capital. The assignment of skills to tasks is endogenously based on comparative advantage and routine-biased technical change involves the substitution of machines for certain tasks previously performed by workers. We add a city-specific absolute and comparative advantage in producing tasks in the spirit of Davis and Dingel (2020) and Davis et al. (2023). We integrate this setting for the demand side of the economy into a standard spatial equilibrium framework, where heterogeneous workers sort across cities based on endogenous city characteristics. The absolute and comparative advantage of cities and workers play a central role in driving wages, city sizes, and sorting of workers across tasks and cities in spatial equilibrium. Overall, we find that routine-biased technical change leads to an increase in spatial wage inequality and skill sorting, consistent with our stylized facts.

Related Literature. This paper is related to three main strands of literature. First, our paper is related to the work on the evolution of spatial wage inequality. Baum-Snow and Pavan (2013) and Baum-Snow et al. (2018) document that wage inequality has increased more in large cities in the US between 1980 and 2007, indicating that the drivers behind wage inequality have an urban bias. Moreover, spatial wage inequality is accompanied by a marked skill bias of agglomeration economies (Diamond, 2016; Grujovic, 2021; Card et al., 2023, among others). Though not only is the urban wage gradient skill-biased, but the gradients have also evolved in opposite directions over the last several decades: the urban wage gradient for college workers has risen, whereas it fell for non-college workers (Berry and Glaeser, 2005; Baum-Snow et al., 2018; Autor, 2019; Diamond and Gaubert, 2022; Butts et al., 2023)[1]. To explain this diverging pattern, Baum-Snow et al. (2018) employ a structural model that incorporates capital skill-complementarity and agglomeration forces. They identify increasing complementarity between cities and skills over time, i.e., rising agglomeration forces for skilled workers, as the main driver. Giannone (2022) also shows with a structural model that wage divergence among skilled workers across cities is driven by rising agglomeration forces for skilled workers that interact positively with skill-biased technical change. Eeckhout et al. (2022) highlight that the more intensive adoption of IT in expensive locations can explain, at least in part, the skill bias in agglomeration economies.

Second, our study is connected to the literature on spatial sorting of heterogeneous agents, recently reviewed in Diamond and Gaubert, 2022. Moretti (2004), Berry and Glaeser (2005), and Diamond (2016) document that the skill ratio of workers was diverging across cities in the US between 1980 and 2010. Moretti (2012) labels this increase in the spatial sorting by skill with the term 'The Great Divergence'. Among the drivers of spatial skill sorting, our work is most closely related to the literature that puts forward changes in labor demand as the force behind the increase in skill sorting (e.g., Behrens and Robert-Nicoud, 2015; Eeckhout et al., 2014; Behrens et al., 2014).

Third, we build on a large literature on labor market polarization and the disappearing jobs at the middle of the wage distribution, reviewed in Acemoglu and Autor (2011). Workers in the middle of the wage distribution are concentrated in manufacturing occupations with a high content of routine tasks, whereas manual service jobs are located at the bottom of the

[1]Butts et al. (2023) find that the urban wage premium declined for college and non-college workers in the U.S. over the period 1940-2010. However, since 1980 the urban wage premium for college-educated workers has started to rise again, whereas the urban wage premium for non-college workers continues to decline.

distribution and abstract professional and managerial jobs at the top (e.g., Autor et al., 2006, and Goos and Manning, 2007). As the main drivers of polarization, much of the focus of this literature is on rapid improvements in the productivity of information and communication technologies affecting routine jobs (e.g., Autor et al., 2003) as well as on international trade and offshoring affecting manufacturing jobs (e.g., Autor et al., 2013).[2]

Further, there is a pronounced geographical dimension to labor market polarization (e.g., Autor and Dorn, 2013), with strong variation in polarization across local labor markets. In a recent paper, Autor (2019) finds that the employment share of middle-paid occupations declined more sharply in denser areas. In particular, polarization has meant a profound reallocation of non-college workers in urban areas from middle-paid production and office jobs to low-paid service occupations. The recent literature focuses on the use of quantitative spatial equilibrium models to rationalize the spatial dimension of polarization. As explanations, faster skill-biased technological change in larger cities through consumption spillovers (Cerina et al., 2023), higher IT adoption in larger cities due to a higher costs of living and higher wages (Eeckhout et al., 2022), the interaction of routinization or offshoring shocks and the comparative advantage of cities (Davis et al., 2023), and housing affordability for middle-income households (Parkhomenko, 2022) have been put forward.

The remainder of the paper is organized as follows. Section 2 describes the data and the stylized facts on spatial wage inequality, skill sorting, and employment polarization. Section 3 develops a spatial equilibrium model that relates employment polarization to increasing spatial wage inequality and skill sorting. Section 4 concludes.

2 Stylized Facts

In this section, we document three key facts on spatial wage inequality, skill sorting, and employment polarization based on German data for the period 1980-2010. First, the gap between the average wages of high and low skill workers has become more positively related with city density since 1980, i.e., spatial wage inequality has increased over time. This increase in spatial wage inequality is accompanied by an increase in the urban wage gradient for high skill workers, whereas the gradient for low skill workers declined, indicating a rise in the skill bias of agglomeration economies. Second, the skill ratio of high to low skill workers has diverged across cities, with denser cities becoming more high skilled (i.e., increased spatial skill sorting). Third, there is a spatial dimension to employment polarization, with

[2] As studied in, e.g., Spitz-Oener (2006), Dustmann et al. (2009), Goos et al. (2014), and Böhm et al. (2022), Germany is no exception to this phenomenon. Dauth (2014) documents that employment polarization in Germany almost exclusively occurs in urban areas, whereas rural regions are barely affected.

polarization being more pronounced in denser cities and affecting especially low skill workers.

2.1 Data

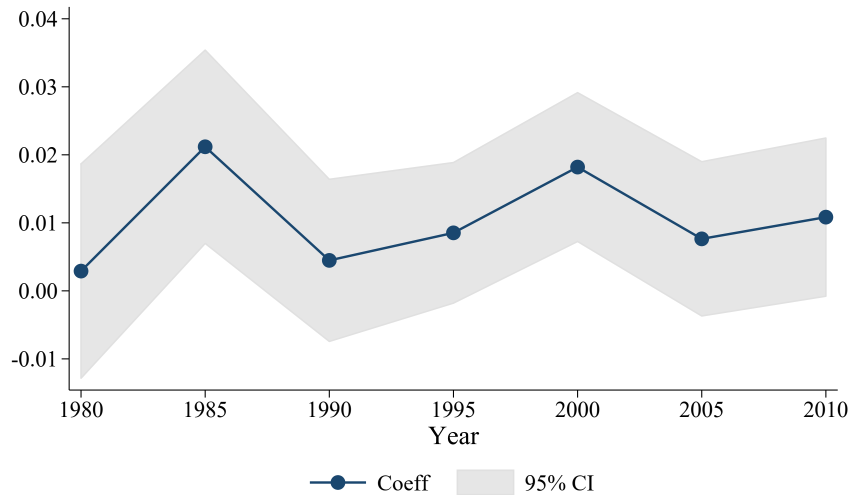
We use the Regional File of the Sample of Integrated Labor Market Biographies (SIAB-R) provided by the Institute for Employment Research (IAB) of the German Federal Employment Agency (Antoni et al., 2019). This panel dataset provides a 2% random sample of administrative social security records and covers employees subject to social security contributions and marginal part-time employment. We restrict the sample to the period between 1980 and 2010, to West Germany, and men in full-time employment liable to social security between the age of 18 and 62. Our definition of skill groups uses education as a proxy for skill: high skill workers are the ones with a university or applied university degree and low skill workers are those with less than that. We aggregate the 120 occupations in the data into four broader groups following the classification of Böhm et al. (2022): managers, professionals, and technicians; sales and office workers; production workers, operators and craftsmen; and workers in services and care occupations. The workplace of an individual is observable at the level of district regions which is our geographical unit of analysis. We will be referring to these district regions as cities. Further details on the data and the definition of variables are provided in appendix B.1.

2.2 Spatial Wage Inequality

We define wage inequality as the difference between the wages of the high skill workers and low skill workers. To document the level of and change in spatial wage inequality from 1980 to 2010, we estimate the following set of bivariate regressions: For each year between 1980 and 2010, we regress logarithmized relative wage of high skill to low skill workers in a city on the logarithmized population density of the city. The estimated coefficients and the corresponding 95% confidence intervals are plotted in figure 1 for every fifth year.

In 1980, the estimated coefficient of log city density on the relative wage of high to low skill workers in a city is somewhat positive, but statistically not significant. This implies that wage inequality between high and low skill workers was not related to city density in 1980 and relatively even distributed across cities. Over the next decades, the estimated coefficient increased to a statistically significant estimate of about 0.01 in 2010. Hence, a positive relationship between the relative wage of high skill to low skill workers in a city and the population density of the city has built up between 1980 and 2010. Denser cities are characterized by higher relative wages of high skill to low skill workers in 2010 and spatial wage inequality between skill groups has emerged between 1980 and 2010.

Figure 1: Spatial wage inequality between high and low skill workers, 1980-2010



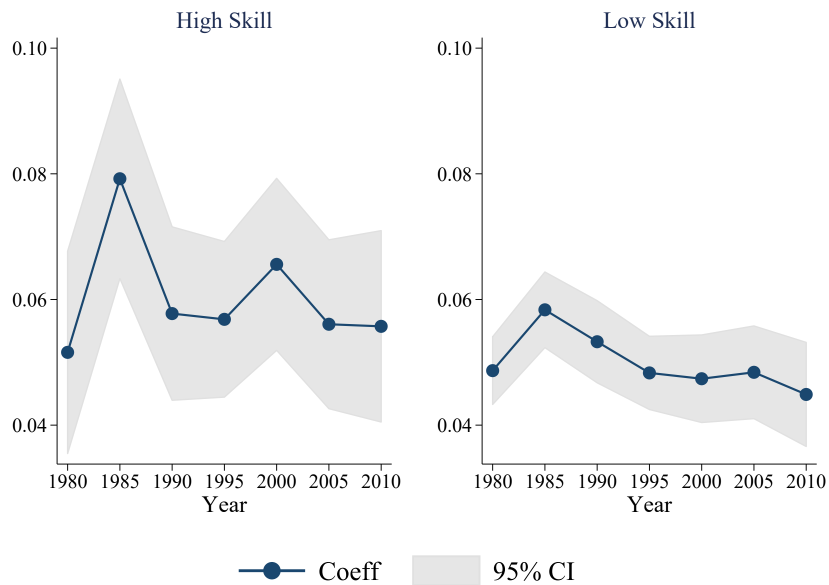
Notes: The figure plots the estimated coefficient and the corresponding 95% confidence interval from a bivariate regression of the logarithm of relative wages of high skill to low skill in a city on the logarithmized population density of the city. The regression is run for the years 1980, 1985, 1990, 1995, 2000, 2005, and 2010, separately. See section B.1 for further details on the data.

To better understand why spatial wage inequality has emerged, we estimate the another set of bivariate regressions: For each year between 1980 and 2010, we regress average logarithmized wage of high skill in a city on the logarithmized population density of the city. The estimated coefficients and the corresponding 95% confidence intervals are plotted in panel A of figure 2 for every fifth year. We then repeat the regression for the average logarithmized wage of low skill workers as the dependent variable. The estimated coefficients and the corresponding 95% confidence intervals are plotted in panel B of figure 2.

In 1980, the estimated coefficient of log city density on the wage of high skill and low skill workers in a city is positive and statistically significant in both cases. This implies that the wage of high and low skill workers is increasing in city density, i.e., there is an urban wage gradient for both skill groups. However, the estimated coefficient for high skill workers is larger than for low skill workers, indicating that agglomeration economies are skill-biased towards high skill workers. Over the next decades, the urban wage gradient of the two skill groups evolved in opposite directions. The estimated coefficient increased slightly for high skill workers, whereas the coefficient decreased for low skill workers. Hence, the urban wage gradient has risen for high skill workers and has fallen for low skill workers between 1980 and 2010. This points at rising agglomeration economies for high skill workers, but decreasing agglomeration economies for low skill workers, i.e., an increase in the skill bias of agglomeration economies.

In summary, the gap between the average wages of high and low skill workers has become

Figure 2: Wages across cities by skill, 1980-2010



Notes: The figure plots the estimated coefficient and the corresponding 95% confidence interval from a bivariate regression of the logarithm of wages for high skill workers (panel A) and low skill workers (panel B) in a city on the logarithmized population density of the city. The regression is run for the years 1980, 1985, 1990, 1995, 2000, 2005, and 2010, separately. See section B.1 for further details on the data.

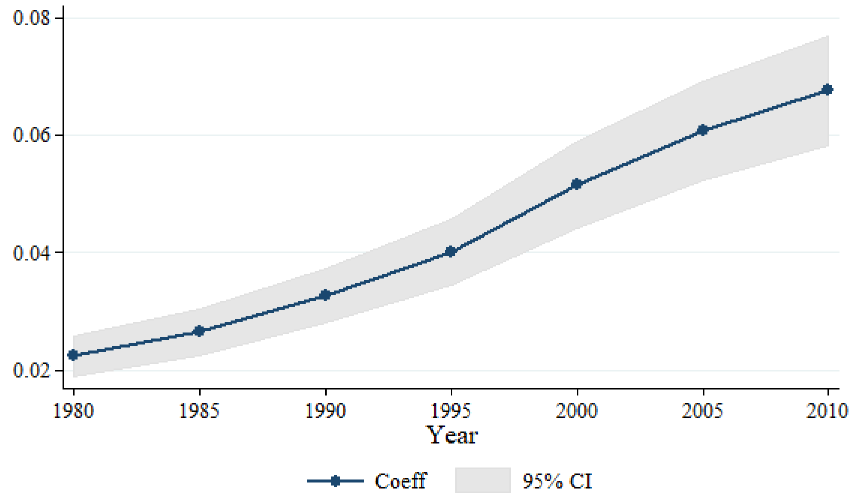
more positively related with city density since 1980, i.e., spatial wage inequality has increased over time. This increase in spatial wage inequality is accompanied by an increase in the urban wage gradient for high skill workers, whereas the gradient for low skill workers declined, indicating a rise in the skill-bias of agglomeration economies.

2.3 Spatial Skill Sorting

We define spatial skill sorting as the differential location decision of high and low skill workers across cities. To document the level of and change in spatial skill sorting from 1980 to 2010, we estimate the following set of bivariate regressions: For each year between 1980 and 2010, we regress the share of high skill workers to low skill workers in a city on the logarithmized population density of the city. The estimated coefficients and the corresponding 95% confidence intervals are plotted in figure 3 for every fifth year.

In 1980, the estimated coefficient of log city density on the ratio of high to low skill workers in a city is positive, around 0.02, and statistically significant. An increase in the population density of a city by 100% is associated with an increase in the share of high skill workers to low skill workers by 2 percentage points. This implies that already in 1980 high skill workers have differentially sorted into denser cities whereas low skill workers have sorted

Figure 3: Spatial sorting by skill, 1980-2010



Notes: The figure plots the estimated coefficient and the corresponding 95% confidence interval from a bivariate regression of the share of high skill workers to low skill workers in a city on the logarithmized population density of the city. The regression is run for the years 1980, 1985, 1990, 1995, 2000, 2005, and 2010, separately. See section B.1 for further details on the data.

into less-dense cities. Over the next decades, the estimated coefficient has steadily risen to an estimate of about 0.07 in 2010 implying a more than threefold increase since 1980. Hence, the positive relationship between the share of high skill to low skill workers in a city and the population density of the city has become stronger between 1980 and 2010. Denser cities are characterized by an even high share of high skill workers in 2010 than in 1980 and spatial skill sorting has been increasing.

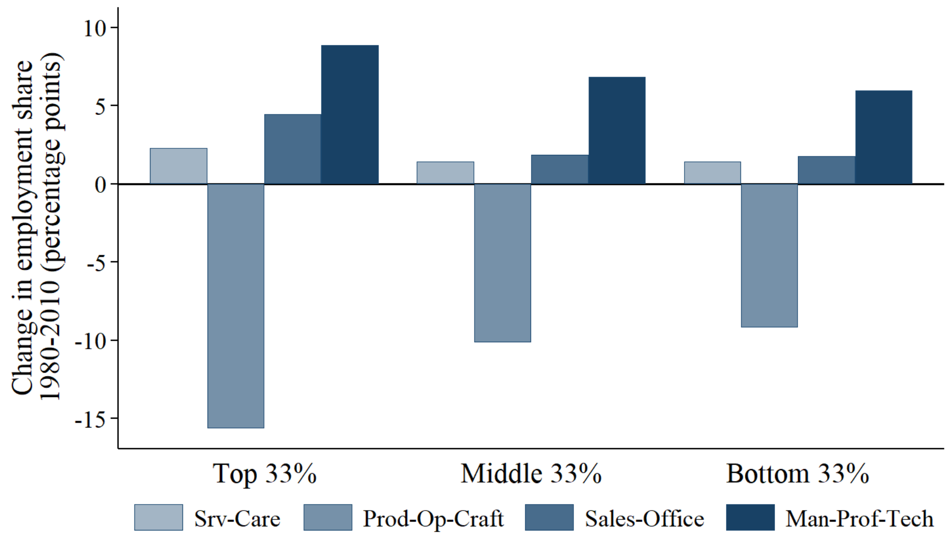
In sum, we find that the skill ratio of high to low skill workers has diverged across cities with denser cities becoming more high skilled, i.e., spatial skill sorting has increased over the time period 1980-2010.

2.4 The Geography of Employment Polarization

Figure 4 depicts the percentage point change in employment shares across cities in West Germany for the period 1980-2010. Occupations are grouped into four categories: managers, professionals, and technicians (Mgr-Prof-Tech); sales and office workers (Sales-Office); production workers, operators, and craftsmen (Prod-Op-Crafts); service and care workers (Srvc-Care). Cities are split into three groups by population density in 2011 such that each group includes one third of all cities.

In the aggregate, figure 4 shows the familiar trend of employment polarization and the disappearance of middle-paying occupations in the West German labor market. The share

Figure 4: Spatial dimension of employment polarization, 1980-2010



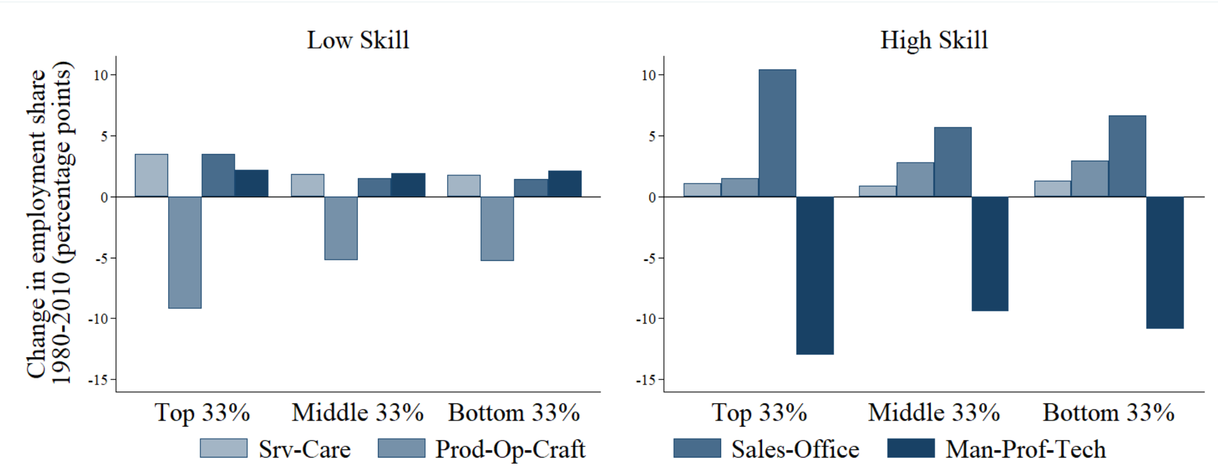
Notes: The figure shows the percentage point change in employment shares across cities in West Germany between 1980 and 2010. The sample is split into three groups of cities by population density in 2011 such that each group includes an equal number of cities. Occupation groups are defined according to the classification of Böhm et al. (2022). See section B.1 for further details on the data.

of employment in Prod-Op-Crafts occupations, that tend to be in the middle of the wage distribution, declined between 1980 and 2010 from 63.4 to 51.3 percent. This fall was offset by an increase in the share of employment in Mgr-Prof-Tech, Sales-Office, and Srvc-Care occupations. Employment in Mgr-Prof-Tech and Sales-Office occupations, which are located at the top of the wage distribution, grew from 16.9 to 24.2 percent and from 15.3 to 18.3 percent, respectively. The share of employment in Service-Care occupations at the bottom of the wage distribution grew more moderately from 4.4 to 6.3 percent. Thus, employment polarization meant that a large share of employment has shifted from middle-paying jobs to high-paying jobs, and only a small share has shifted to low-paying jobs.

Moreover, figure 4 shows that there is a marked geographical dimension to employment polarization. Even though denser cities have historically been less intensive in Prod-Op-Crafts occupations, the decline of these occupation over the last decades was stronger in denser cities. The stronger decline in Prod-Op-Crafts occupations was offset by a higher increase in the employment share in Mgr-Prof-Tech and Sales-Office occupations as well as Srvc-Care occupations. As a result, the positive relationship between the employment share in these occupations and the population density of a city has increased over the last decades. Hence, there is a spatial dimension to employment polarization, with polarization having been more pronounced in cities with a higher population density between 1980 and 2010.

As workers with different skill levels are not evenly distributed across different occu-

Figure 5: Spatial dimension of employment polarization by skill, 1980-2010



Notes: The figure shows the percentage point change in employment shares for low skill workers (panel A) and for high skill workers (panel B) across cities in West Germany between 1980 and 2010. The sample is split into three groups of cities by population density in 2011 such that each group includes an equal number of cities. Occupation groups are defined according to the classification of Böhm et al. (2022). See section B.1 for further details on the data.

pations, employment polarization potentially affects skill groups heterogeneously. Figure 5 shows percentage point change in employment shares across cities for low skill workers (panel A) and high skill workers (panel B) separately. Low skill workers were predominantly employed in Prod-Op-Crafts occupations in 1980. Between 1980 and 2010, the overall share of low skill workers decreased, but strongest within Prod-Op-Crafts occupations. This decline was offset by a roughly equal increase in the employment share in the other occupation groups though the decline in Prod-Op-Crafts occupations was stronger in denser cities. Thus, the pattern of spatial employment polarization with a decline of Prod-Op-Crafts occupations observed in figure 4 especially affects low skill workers. In 1980, over 80 percent of high skill workers were employed in Mgr-Prof-Tech occupations. Over the ensuing decades, the share of high skill workers increased across all occupation groups but the strongest within Sales-Office occupations such that the share in Mgr-Prof-Tech occupations fell while the share in Sales-Office occupations increased. This occupational reallocation of high skill workers is more pronounced in cities with a higher population density. Hence, high skill workers are partly taking over Sales-Office occupations that were typically performed by low skill workers.

In summary, there is a spatial dimension to employment polarization with polarization being more pronounced in denser cities and affecting especially low skill workers.

3 Theoretical Framework

To highlight the role of occupational change for the increase in spatial wage inequality and skill sorting, we build a model that integrates the task-based framework in Acemoglu and Autor (2011) with the spatial equilibrium model of Diamond and Gaubert (2022). As in Acemoglu and Autor (2011), the assignment of skills to tasks is endogenously based on comparative advantage and routine-biased technical change involves the substitution of machines for certain tasks previously performed by workers. To this setting, we add an absolute and comparative advantage of cities in producing tasks. As in Diamond and Gaubert (2022), heterogeneous workers sort across cities based on endogenous city characteristics. The absolute and comparative advantage of cities and workers in producing tasks will then play a central role for driving wages, city sizes and sorting of workers across tasks and cities in spatial equilibrium. Overall, we find that routine-biased technical change leads to an increase in spatial wage inequality and skill sorting, consistent with the stylized facts presented in the previous section.

Environment. We consider a static economy with a discrete set of cities $i \in \mathcal{I}$. There are two types s of workers, high skill h and low skill l, where the aggregate supply of each skill group L^s is exogenously given. Workers consume a single final good and housing, have idiosyncratic preferences for cities, and freely choose a city for living. Each worker supplies inelastically one unit of labor. The single final good is produced in each city by combining a continuum of different tasks. Tasks are produced with high and low skill workers as well as capital. The production of tasks is characterized by an absolute and comparative advantage of cities and workers in producing different tasks. Both, the final good and tasks, are produced under perfect competition. In our baseline model, tasks are only produced with heterogeneous labor. We then extend the framework in section 3.5 to study the role of routine-biased technical change.

3.1 Local Labor Demand

Production. There is a unique final good that is produced in each city by combining a continuum of tasks indexed by $\sigma \in [0, 1]$. Output of the final good is given by the following CES production function

$$Y_i = \left[\int_0^1 q_i(\sigma)^{\frac{\epsilon-1}{\epsilon}} d\sigma\right]^{\frac{\epsilon}{\epsilon-1}}, \tag{1}$$

where $q_i(\sigma)$ denotes the production level of task σ in city i and $\epsilon > 0$ is the elasticity of substitution across tasks. The final good is freely traded across cities and its price is set as the numéraire.

Tasks are non-tradable and produced with high and low skill workers indexed by $s \in \{l, h\}$. Workers are perfect substitutes in the production of each task but differ in their productivity. Each task has the following production function

$$q_i(\sigma) = \sum_{s \in \{l,h\}} A_i(\sigma) \alpha^s(\sigma) L_i^s(\sigma), \tag{2}$$

where $L_i^s(\sigma)$ denotes the endogenous number of workers with skill s producing task σ in city i, $A_i(\sigma)$ is the city-specific productivity for producing task σ, and $\alpha^s(\sigma)$ is the skill-specific productivity for producing task σ.

The skill-specific productivity $\alpha^s(\sigma)$ reflects the absolute and comparative advantage of workers in producing different tasks. We assume that high skill workers have an absolute advantage in performing all tasks, but a comparative advantage in performing higher-σ tasks

$$\alpha^h(\sigma) > \alpha^l(\sigma), \qquad \text{for all } \sigma \in [0, 1]$$

$$\alpha^h(\sigma)\alpha^l(\sigma') > \alpha^h(\sigma')\alpha^l(\sigma) \qquad \text{for all } \sigma > \sigma'.$$

This can be interpreted as stating that higher-σ indices correspond to more complex or abstract tasks in which high skill workers are more productive than low skill workers.

The city-specific productivity $A_i(\sigma)$ captures the absolute and comparative advantage of cities in producing different tasks. We assume that $A_i(\sigma)$ is twice-differentiable, strictly increasing in i, and strictly log-supermodular in i and σ

$$A_i(\sigma) > A_{i'}(\sigma) \qquad \text{for all } i > i' \text{ and } \sigma \in [0, 1]$$

$$A_i(\sigma)A_{i'}(\sigma') > A_i(\sigma')A_{i'}(\sigma) \qquad \text{for all } i > i' \text{ and } \sigma > \sigma'.$$

The former implies that higher-i cities are more productive across all tasks so that cities are indexed by their total factor productivity. The latter states that higher-i cities have a comparative advantage in the production of higher-σ tasks, i.e. more complex tasks. Note that there is a complementary in the comparative advantage of high skill workers and higher-i cities in performing higher-σ tasks, i.e., more complex tasks. These differences in absolute and comparative advantage of workers and cities will play a central role in our model.

Allocation of skills to tasks. We first characterize the allocation of skills to tasks. For each city i, there exists a threshold $\bar{\sigma}_i \in [0, 1]$ such that all tasks $\sigma < \bar{\sigma}_i$ are produced using low skill workers (i.e., $L_i^h(\sigma) = 0$ for all $\sigma < \bar{\sigma}_i$). All tasks $\sigma > \bar{\sigma}_i$ are produced using high skill workers (i.e., $L_i^l(\sigma) = 0$ for all $\sigma > \bar{\sigma}_i$). Hence, the task threshold partitions the continuum of tasks into two distinct sets according to the comparative advantage of workers: the first set including the less complex tasks is performed by low skill workers and the second set with the more complex tasks is performed by high skill workers.[3]

The task threshold is endogenously determined in equilibrium by a 'no-arbitrage' condition that equalizes the cost of producing this threshold task using low or high skill workers. In general, workers of the same skill group receive the same city-specific wage in equilibrium even though they perform different tasks because no worker would supply their labor to tasks paying lower wages. Then, the threshold task $\bar{\sigma}_i$ must be such that this task can be produced with the same costs using either high or low skill workers. This is equivalent to task $\bar{\sigma}_i$ having the same equilibrium supply either when produced only with high or low skill workers. This implies the following no arbitrage condition between high and low skill workers for city i

$$\frac{L_i^h}{L_i^l} = \left(\frac{\alpha^h(\bar{\sigma}_i)}{\alpha^l(\bar{\sigma}_i)}\right)^{-\epsilon} \left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)}\right)^{\epsilon-1}, \tag{3}$$

where $\Omega_i^l(\bar{\sigma}_i) \equiv [\int_0^{\bar{\sigma}_i}(A_i(\sigma')\alpha^l(\sigma'))^{\epsilon-1}d\sigma']^{\frac{1}{\epsilon-1}}$ and $\Omega_i^h(\bar{\sigma}_i) \equiv [\int_{\bar{\sigma}_i}^1(A_i(\sigma')\alpha^h(\sigma'))^{\epsilon-1}d\sigma']^{\frac{1}{\epsilon-1}}$. The task threshold is determined by the skill-specific and the city-specific productivity for producing tasks as well as the equilibrium ratio for high to low skill workers in a city.

Local labor demand. Since the labor market is perfectly competitive, wages equal the marginal product of labor for high and low skill workers, respectively. Using, the no-arbitrage condition for the task threshold, the relative local labor demand curve is

$$\frac{L_i^h}{L_i^l} = \left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)}\right)^{\epsilon-1} \left(\frac{w_i^h}{w_i^l}\right)^{-\epsilon}. \tag{4}$$

Each city faces a downward sloping local labor demand curve because firms are willing to hire more workers as the wage declines. Moreover, local labor demand depends on the task threshold which is endogenously determined in equilibrium.

3.2 Local Labor Supply

Preferences. Workers choose where to live among cities and have Cobb-Douglas preferences for the consumption good and housing with price r_i in city i. Each worker supplies inelastically one unit of labor for wage w_i^s in city i and has idiosyncratic preferences for locations.

[3] For a formal proof, see Acemoglu and Zilibotti (2001) and Acemoglu and Autor (2011).

The utility of worker s, who is of type ω and lives in city i, is

$$\max_{c,h} U_i^s(\omega) = \frac{1}{\beta^\beta (1-\beta)^{1-\beta}} c^{1-\beta} h^\beta \epsilon_i^s(\omega) \qquad \text{such that} \qquad c + r_i h = w_i^s, \tag{5}$$

where c denotes consumption, h denotes housing, and $\epsilon_i^s(\omega)$ is the idiosyncratic preference shock for living in city i. This idiosyncratic preference shock is independent and identically Fréchet distributed across workers within a skill group and across cities with scale parameter $\kappa > 1$. A larger preference shock $\epsilon_i^s(\omega)$ means that a worker is particularly attached to city i, holding real wages constant.

The worker's maximized utility function, subject to the budget constraint, can be expressed as an indirect utility function for living in city i as

$$V_i^s(\omega) = V_i^s \epsilon_i^s(\omega) = \frac{w_i^s}{r_i^\beta} \epsilon_i^s(\omega), \tag{6}$$

where $V_i^s \equiv w_i^s / r_i^\beta$ is the systematic indirect utility component.

Local labor supply. Under the assumption of Fréchet distributed idiosyncratic preferences, the location choice of skill group s can be summarized with the share of workers of skill s who choose city i

$$\frac{L_i^s}{L^s} = \frac{\left(\frac{w_i^s}{r_i^\beta}\right)^\kappa}{\sum_{i' \in \mathcal{I}} \left(\frac{w_{i'}^s}{r_{i'}^\beta}\right)^\kappa} = \frac{V_i^{s\kappa}}{\sum_{i' \in \mathcal{I}} V_{i'}^{s\kappa}} = \frac{V_i^{s\kappa}}{V^{s\kappa}}, \tag{7}$$

where L_i^s is the number of workers with skill s who choose city i and $V^s \equiv \left(\sum_{j \in \mathcal{I}} V_j^{s\kappa}\right)^{\frac{1}{\kappa}}$ is the expected utility for a worker with skill s across all cities. The parameter κ captures the elasticity of population shares with respect to real wages and is therefore a measure of mobility of skill group s.

The relative local labor supply of high skill to low skill workers in city i is then given by

$$\frac{L_i^h}{L_i^l} = \left(\frac{w_i^h}{w_i^l}\right)^\kappa \left(\frac{V^h}{V^l}\right)^{-\kappa} \left(\frac{L^h}{L^l}\right). \tag{8}$$

Each city faces an upward sloping local labor supply curve because workers are attracted by higher wages. As high and low skill workers have the same housing expenditure share, the relative local labor supply only depends on the relative wage and not on the housing price.

3.3 Housing Supply

For simplicity, we assume that housing is owned by absentee landlords who supply the housing to local residents. The local housing price in city i is given by

$$r_i = \bar{r} L_i^{\gamma}, \tag{9}$$

where L_i denotes the population of city i, γ is the (inverse) housing supply elasticity and $\bar{r}$ the part of the local housing price that is independent of the size of the city. The housing supply elasticity captures, in a reduced-form way, forces that hinder the expansion of the housing stock such as geographical constraints and land-use regulations.

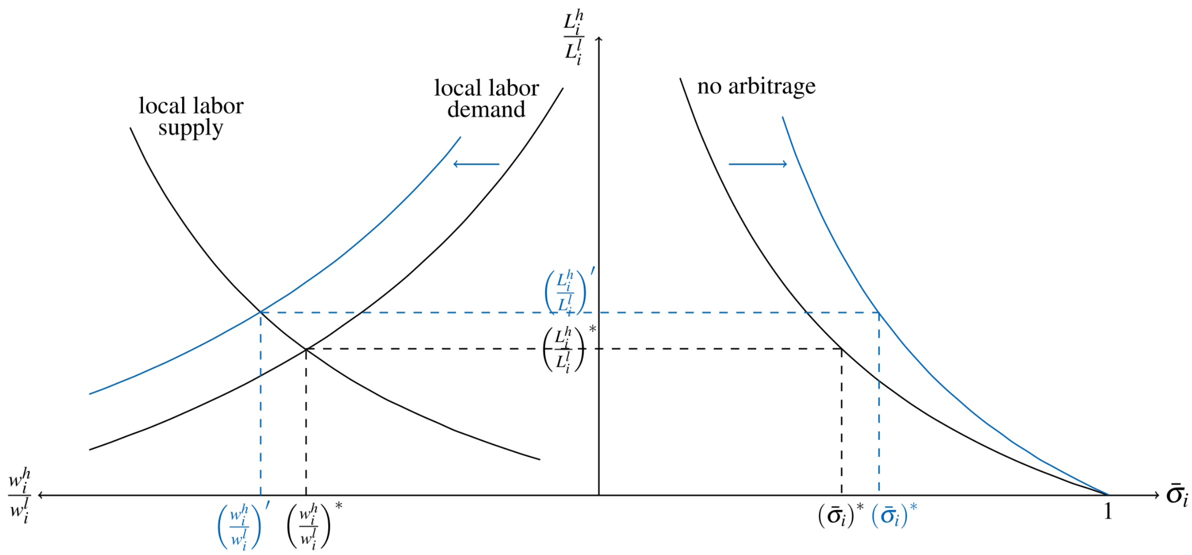
3.4 Spatial Equilibrium

Definition of a spatial equilibrium. In a competitive spatial equilibrium, individuals maximize their utility, final goods and intermediate tasks producers maximize profits, and markets clear. Given the city-specific productivity in each city ($A_i(\sigma)$) and the skill-specific productivity for each skill group ($\alpha^s(\sigma)$), the equilibrium is defined by a menu of task thresholds ($\bar{\sigma}_i^*$), wages (w_i^{l*}, w_i^{h*}), and rents (${r_i}^*$) with population levels (L_i^{l*}, L_i^{h*}) such that equations (3), (4), (8), and (9) are satisfied for all cities indexed by i.

Properties of the spatial equilibrium. Solving numerically for the spatial equilibrium, we find that the city size, housing rents, and wages of high and low skill workers, the task threshold, the relative number of high skill workers, and the relative wage of high skill workers are increasing in the total factor productivity of a city, indexed by i. As a result, wage inequality and skill sorting is increasing in city sizes, as documented in sections 2.2 and 2.3 on the stylized facts.

The spatial equilibrium is illustrated in figure 6. This figure shows the relative local labor supply curve from equation (8), the relative local labor demand curve from equation (4), and the 'no-arbitrage' condition pinning down the task threshold from equation (3) for two different cities. City 2, depicted in blue, has a higher absolute advantage (higher total factor productivity) in producing all tasks than city 1, depicted in black. As city 2 has a higher absolute advantage and comparative advantage in producing more complex tasks than city 1, the local labor demand curve and the curve for the 'no-arbitrage' condition of city 2 lie further out compared to city 1. Both cities share the same local labor supply curve. This implies that the equilibrium task threshold, relative number of high skill workers, and relative wage of high skill workers in city 2 are higher than in city 1.

Figure 6: Illustration of the spatial equilibrium



Notes: The figure illustrates the local labor supply curve, the local labor demand curve, and the 'no-arbitrage' condition for two different cities. City 2, depicted in blue, has a higher absolute advantage (higher total factor productivity) in producing all tasks than city 1, depicted in black, i.e., $A_2(\sigma) > A_1(\sigma)$ for all $\sigma \in [0, 1]$. Both cities have the same labor supply curve, but the labor demand curve and the curve for the 'no-arbitrage' condition lie further out compared to city 1. This implies that the equilibrium task threshold, relative number of high skill workers, and relative wage of high skill workers in city 2 are higher than in city 1.

3.5 Task replacing technologies

Capital. We now use our model to investigate the implications of routine-biased technical change that involves the substitution of capital (machines) for certain tasks previously performed by workers. In general, tasks performed by all two skill groups are subject to machine displacement, but based on the empirical evidence documented in section 2.4, the set of tasks most subject to machine displacement are those in production occupations, primarily, though not exclusively, performed by low skill workers.

Theoretically, we assume that tasks can now be produced with high and low skill workers as well as capital. Thus, the production function of tasks is adjusted as follows

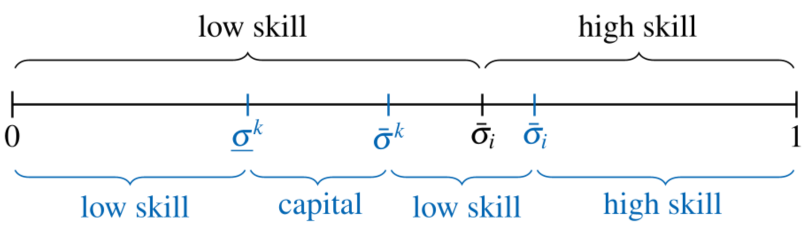
$$q_i(\sigma) = \sum_{s\in\{l,h\}} A_i(\sigma)\alpha^s(\sigma)L_i^s(\sigma) + \alpha^k(\sigma)K_i(\sigma), \tag{10}$$

where $K_i(\sigma)$ denotes the amount for capital used for producing task σ in city i, and $\alpha^k(\sigma)$ is the capital-specific productivity for producing task σ. The productivity of capital is assumed to be the same across cities and hence has no inherent bias towards specific cities. The absolute and comparative advantage of cities and workers is specified as in section 3.1.

We assume that the capital-specific productivity $\alpha^k(\sigma)$ is large enough (with a fixed cost of capital r) such that there exists a range of tasks $[\underline{\sigma}^k, \bar{\sigma}^k] \subset [0, \bar{\sigma}_i]$ that are now performed with lower costs by machines than low skill workers. For the remaining tasks, i.e., for all $\sigma \notin [\underline{\sigma}^k, \bar{\sigma}^k]$, we continue to assume that $\alpha^k(\sigma) = 0$. The range of tasks that is now performed by capital is for simplicity not city-specific.

New allocation of skills to tasks. After the introduction of capital, there exists a new equilibrium characterized by a new threshold $\bar{\sigma}_i'$ such that $0 < \underline{\sigma}^k < \bar{\sigma}^k < \bar{\sigma}_i' < 1$. All tasks $\sigma \in \left(0, \underline{\sigma}^k\right) \cup \left(\bar{\sigma}^k, \bar{\sigma}_i'\right)$ are performed by low skill workers (i.e., $L_i^h(\sigma) = 0$), and all tasks $\sigma > \bar{\sigma}_i'$ are performed by high skill workers (i.e., $L_i^l(\sigma) = 0$) in city i. Hence, there is a reallocation of tasks as a consequence of capital replacing tasks previously performed by low skill workers. In particular, low skill workers will now start performing some of the tasks previously allocated to high skill workers (corresponding to sales and office occupations in section 2.4 and expanding their supply of the least complex tasks (corresponding to service and care occupations in section 2.4). High skill workers instead lose the, from their perspective, least complex tasks to low skill workers and expand their supply of the most complex tasks. The reallocation is illustrated in figure 7.

Figure 7: New allocation of skills to tasks under routine-biased technological change



Notes: The figure illustrates the allocation of skills to tasks. The allocation without capital is depicted in black and the new allocation with capital is shown in blue. There is a reallocation of tasks as a consequence of capital replacing tasks previously performed by low skill workers. In particular, low skill workers will now start performing some of the tasks previously allocated to high skill workers and expanding their supply of the least complex tasks. High skill workers instead lose their least complex tasks to low skill workers and expand their supply of the most complex tasks.

Because new machines replace the tasks previously performed by low skill workers, their relative wage compared to high skill workers declines. As high skill workers expand their supply of the most complex tasks where also higher-i cities have a comparative advantage, the relative wage of high skill workers increases stronger in higher-i cities whereby higher-i cities are larger in equilibrium. Hence, spatial wage inequality increases which explains the first stylized fact from section 2.2. Attracted by the stronger increase in relative wages of high skill workers in higher-i cities, high skill workers move to these cities. As a result, housing rents are increasing as well, which prices out low skill workers in these cities. Instead, low skill workers are moving to lower-i cities which are smaller in equilibrium. Thus, the share of high skill workers rises in the largest cities and falls in the smallest cities. This explains the second stylized from section 2.3 referring to the increase in spatial skill sorting.

4 Conclusion

In this paper, we establish three stylized facts on the evolution of spatial wage inequality, skill sorting, and employment polarization in Germany between 1980 and 2010. First, we find that the gap between the average wages of high and low skill workers has become more positively related with city density since 1980, i.e., spatial wage inequality has increased over time. This increase in spatial wage inequality is accompanied by an increase in the urban wage gradient for high skill workers, whereas the gradient for low skill workers declined, indicating a rise in the skill-bias of agglomeration economies. The second fact is that the skill ratio of high to low skill workers has diverged across cities with denser cities becoming more high skilled (i.e., increased spatial skill sorting). Finally, we show that there is a spatial dimension to employment polarization with polarization being more pronounced in denser cities and affecting especially low skill workers.

Motivated by these facts, we develop a spatial equilibrium model to highlight the role of occupational change for the increase in spatial wage inequality and skill sorting. Building on the literature on the task-based framework, we consider a continuum of tasks that is produced with heterogeneously skilled workers and capital. The assignment of skills to tasks is endogenously based on comparative advantage and routine-biased technical change involves the substitution of machines for certain tasks previously performed by workers. We add a city-specific absolute and comparative advantage in producing tasks. We integrate this setting for the demand side of the economy into a standard spatial equilibrium framework, where heterogeneous workers sort across cities based on endogenous city characteristics. The absolute and comparative advantage of cities and workers play a central role in driving wages, city sizes, and sorting of workers across tasks and cities in spatial equilibrium.

Overall, we find that routine-biased technical change leads to an increase in spatial wage inequality and skill sorting, consistent with our stylized facts.

References

Acemoglu, Daron and David H. Autor (2011): Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4(12), pp. 1043–1171.

Acemoglu, Daron and Fabrizio Zilibotti (2001): Productivity Differences. *Quarterly Journal of Economics*, 116(2), pp. 563–606.

Antoni, Manfred, Andreas Ganzer, and Philipp vom Berge (2019). *Institute for Employment Research, FDZ-Datenreport*, 04/2019.

Autor, David H. (2019): Work of the Past, Work of the Future. *AEA Papers and Proceedings*, 109, pp. 1–32.

Autor, David H. and David Dorn (2013): The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5), pp. 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson (2013): The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6), pp. 2121–2168.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2006): The Polarization of the US Labor Market. *American Economic Review*, 96(2), pp. 189–194.

— (2008): Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2), pp. 300–323.

Autor, David H., Frank Levy, and Richard J. Murnane (2003): The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118(4), pp. 1279–1333.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange (2009): Skills in the city. *Journal of Urban Economics*, 65(2), pp. 136–153.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan (2018): Why Has Urban Inequality Increased? *American Economic Journal: Applied Economics*, 10(4), pp. 1–42.

Baum-Snow, Nathaniel and Ronni Pavan (2013): Inequality and City Size. *Review of Economics and Statistics*, 95(5), pp. 1535–1548.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud (2014): Productive Cities: Sorting, Selection, and Agglomeration. *Journal of Political Economy*, 122(3), pp. 507–553.

Behrens, Kristian and Frédéric Robert-Nicoud (2015): Agglomeration Theory with Heterogeneous Agents. *Handbook of Regional and Urban Economics*, 5(4), pp. 171–245.

Berry, Christopher L. and Edward L. Glaeser (2005): The Divergence of Human Capital Levels across Cities. *Papers in Regional Science*, 84(3), pp. 407–444.

Böhm, Michael J., Hans-Martin von Gaudecker, and Felix Schran (2022): Occupation Growth, Skill Prices, and Wage Inequality. *Accepted: Journal of Labor Economics.*

Butts, Kyle, Taylor Jaworski, and Carl Kitchens (2023): The Urban Wage Premium in Historical Perspective. *NBER Working Paper*, 31387.

Card, David, Jörg Heining, and Patrick Kline (2013): Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics*, 128(3), pp. 967–1015.

Card, David, Jesse Rothstein, and Moises Yi (2023): Location, Location, Location.

Cerina, Fabio, Elisa Dienesch, Alessio Moro, and Michelle Rendall (2023): Spatial Polarisation. *The Economic Journal*, 133(649), pp. 30–69.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008): Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63, pp. 723–742.

Combes, Pierre-Philippe and Laurent Gobillon (2015): The Empirics of Agglomeration Economies. *Handbook of Regional and Urban Economics*, 5, pp. 247–348.

Costinot, Arnaud and Jonathan Vogel (2010): Matching and Inequality in the World Economy. *Journal of Political Economy*, 118 (4), pp. 747–786.

Dauth, Wolfgang (2014): Job Polarization on Local Labor Markets. *IAB-Discussion Paper*, 18/2014.

Dauth, Wolfgang and Johann Eppelsheimer (2020): Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide. *Journal for Labour Market Research*, 54(10).

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2022): Matching in Cities. *Journal of the European Economic Association*, 20(4), pp. 1478–1521.

Davis, Donald R. and Jonathan I. Dingel (2020): The comparative advantage of cities. *Journal of International Economics*, 123, 103291.

Davis, Donald R., Eric Mengus, and Tomasz Michalski (2023): Labor Market Polarization and The Great Urban Divergence. *NBER Working Paper*, 26955.

De la Roca, Jorge and Diego Puga (2017): Learning by working in bigger cities. *Review of Economic Studies*, 84(1), pp. 106–142.

Diamond, Rebecca (2016): The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000. *American Economic Review*, 106(3), pp. 479–524.

Diamond, Rebecca and Cecile Gaubert (2022): Spatial Sorting and Inequality. *Annual Review of Economics*, 14, pp. 795–819.

Duranton, Gilles and Diego Puga (2004): Micro-foundations of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics*, 4(48), pp. 2063–2117.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009): Revisiting the German Wage Structure. *Quarterly Journal of Economics*, 124(2), pp. 843–881.

Eberle, Johanna and Alexandra Schmucker (2019): Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2017 - programming examples for Stata. *Institute for Employment Research, FDZ-Methodenreport*, 04/2019.

Eeckhout, Jan, Christoph Hedtrich, and Roberto Pinheiro (2022): IT and Urban Polarization.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny (2014): Spatial Sorting. *Journal of Political Economy*, 122(3), pp. 554–620.

Giannone, Elisa (2022): Skill-biased Technical Change and Regional Convergence.

Glaeser, Edward L. and David C. Maré (2001): Cities and Skills. *Journal of Labor Economics*, 19(2), pp. 316–342.

Goos, Maarten and Alan Manning (2007): Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *Review of Economics and Statistics*, 89(1), pp. 118–133.

Goos, Maarten, Alan Manning, and Anna Salomons (2014): Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8), pp. 2509–2526.

Grujovic, Anja (2021): Tasks, Cities and Urban Wage Premia.

Koster, Hans R.A. and Ceren Ozgen (2021): Cities and tasks. *Journal of Urban Economics*, 126, 103386.

Moretti, Enrico (2004): Human Capital Externalities in Cities. *Handbook of Regional and Urban Economics*, 4(51), pp. 2243–2291.

— (2012): *The New Geography of Jobs*. Houghton Mifflin Harcourt.

Parkhomenko, Andrii (2022): Homeownership, Polarization, and Inequality.

Spitz-Oener, Alexandra (2006): Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2), pp. 235–270.

Appendix

A Theory Appendix

A.1 Local Labor Demand

Production. There is a unique final good that is produced in each city by combining a continuum of tasks indexed by $\sigma \in [0, 1]$. Output of the final good is given by the following CES production function

$$Y_i = \left[\int_0^1 q_i(\sigma)^{\frac{\epsilon-1}{\epsilon}} d\sigma \right]^{\frac{\epsilon}{\epsilon-1}}, \tag{A.1}$$

where $q_i(\sigma)$ denotes the production level of task σ in city i and $\epsilon > 0$ is the elasticity of substitution across tasks. The final good is freely traded across cities and its price is set as the numéraire.

Let $p_i(\sigma)$ denote the price of task σ in city i. Then, profit maximization of the final good suppliers implies

$$\max_{q_i(\sigma)} \Pi = \left[\int_0^1 q_i(\sigma)^{\frac{\epsilon-1}{\epsilon}} d\sigma \right]^{\frac{\epsilon}{\epsilon-1}} - \int_0^1 p_i(\sigma) q_i(\sigma) d\sigma$$

$$\Rightarrow \quad q_i(\sigma) = p_i(\sigma)^{-\epsilon} Y_i.$$

Tasks are non-tradable and produced with high and low skill workers indexed by $s \in \{l, h\}$. Workers are perfect substitutes in the production of each task but differ in their productivity. Each task has the following production function

$$q_i(\sigma) = \sum_{s \in \{l,h\}} A_i(\sigma) \alpha^s(\sigma) L_i^s(\sigma), \tag{A.2}$$

where $L_i^s(\sigma)$ denotes the endogenous number of workers with skill s producing task σ in city i, $A_i(\sigma)$ is the city-specific productivity for producing task σ, and $\alpha^s(\sigma)$ is the skill-specific productivity for producing task σ.

The skill-specific productivity $\alpha^s(\sigma)$ reflects the absolute and comparative advantage of workers in producing different tasks. We assume that high skill workers have an absolute advantage in performing all tasks, but a comparative advantage in performing higher-σ tasks

$$\alpha^h(\sigma) > \alpha^l(\sigma), \qquad \text{for all } \sigma \in [0, 1]$$

$$\alpha^h(\sigma) \alpha^l(\sigma') > \alpha^h(\sigma') \alpha^l(\sigma) \qquad \text{for all } \sigma > \sigma'.$$

This can be interpreted as stating that higher-σ indices correspond to more complex or abstract tasks in which high skill workers are more productive than low skill workers.

The city-specific productivity $A_i(\sigma)$ captures the absolute and comparative advantage of cities in producing different tasks. We assume that $A_i(\sigma)$ is twice-differentiable, strictly increasing in i, and strictly log-supermodular in i and σ

$$A_i(\sigma) > A_{i'}(\sigma) \qquad \text{for all } i > i' \text{ and } \sigma \in [0,1]$$

$$A_i(\sigma)A_{i'}(\sigma') > A_i(\sigma')A_{i'}(\sigma) \qquad \text{for all } i > i' \text{ and } \sigma > \sigma'.$$

The former implies that higher-i cities are more productive across all tasks so that cities are indexed by their total factor productivity. The latter states that higher-i cities have a comparative advantage in the production of higher-σ tasks, i.e. more complex tasks. Note that there is a complementary in the comparative advantage of high skill workers and higher-i cities in performing higher-σ tasks, i.e., more complex tasks.

Let w_i^s denote the wage payed to a worker with skill s in city i. Then, profit maximization of the suppliers of task σ implies

$$\max_{L_i^s(\sigma)} \pi(\sigma) = \sum_{s \in \{l,h\}} \left[p_i(\sigma)A_i(\sigma)\alpha^s(\sigma) - w_i^s\right] L_i^s(\sigma)$$

$$\Rightarrow \quad w_i^s \geq p_i(\sigma)A_i(\sigma)\alpha^s(\sigma), \forall s \qquad \text{and} \qquad w_i^s = p_i(\sigma)A_i(\sigma)\alpha^s(\sigma), \forall s \text{ with } L_i^s(\sigma) > 0.$$

Allocation of skills to tasks. For each city i, there exists a threshold $\bar{\sigma}_i \in [0,1]$ such that all tasks $\sigma < \bar{\sigma}_i$ are produced using low skill workers (i.e., $L_i^h(\sigma) = 0$ for all $\sigma < \bar{\sigma}_i$). All tasks $\sigma > \bar{\sigma}_i$ are produced using high skill workers (i.e., $L_i^l(\sigma) = 0$ for all $\sigma > \bar{\sigma}_i$).

Even though workers of the same skill level perform different tasks, in equilibrium they receive the same wage for a given city because no worker would supply their labor to tasks paying lower wages. This implies that the value marginal product of all workers in a skill group and city must be the same across all tasks that they are performing

$$w_i^l = p_i(\sigma)A_i(\sigma)\alpha^l(\sigma) \qquad \text{for all } \sigma < \bar{\sigma}_i$$
$$w_i^h = p_i(\sigma)A_i(\sigma)\alpha^h(\sigma) \qquad \text{for all } \sigma > \bar{\sigma}_i.$$

Thus in each city, the price difference between two tasks produced by the same type of worker must offset the productivity difference of this type of worker in these two tasks

$$p_i(\sigma)A_i(\sigma)\alpha^l(\sigma) = p_i(\sigma')A_i(\sigma')\alpha^l(\sigma') \qquad \text{for all } \sigma, \sigma' < \bar{\sigma}_i$$
$$p_i(\sigma)A_i(\sigma)\alpha^h(\sigma) = p_i(\sigma')A_i(\sigma')\alpha^h(\sigma') \qquad \text{for all } \sigma, \sigma' > \bar{\sigma}_i.$$

The first-order condition for profit maximization in the production of the final good and the choice of the final good as the numéraire imply

$$p_i(\sigma)^\epsilon q_i(\sigma) = Y_i.$$

Now consider two tasks σ, σ' performed by workers with skill s. Then using the definition of the productivity of workers with skill s in these tasks, we have

$$p_i(\sigma)^\epsilon A_i(\sigma)\alpha^s(\sigma)L_i^s(\sigma) = p_i(\sigma')^\epsilon A_i(\sigma')\alpha^s(\sigma')L_i^s(\sigma').$$

$$\Rightarrow \qquad L_i^s(\sigma') = \left[\frac{A_i(\sigma')\alpha^s(\sigma')}{A_i(\sigma)\alpha^s(\sigma)}\right]^{\epsilon-1} L_i^s(\sigma)$$

Using the market clearing condition for low skill workers, $L_i^l = \int_0^{\bar{\sigma}_i} L_i^l(\sigma')d\sigma'$, and for high skill workers, $L_i^h = \int_{\bar{\sigma}_i}^1 L_i^h(\sigma')d\sigma'$, we must have

$$L_i^l(\sigma) = \frac{\left[A_i(\sigma)\alpha^l(\sigma)\right]^{\epsilon-1} L_i^l}{\int_0^{\bar{\sigma}_i} \left(A_i(\sigma')\alpha^l(\sigma')\right)^{\epsilon-1} d\sigma'}, \qquad \text{for all } \sigma < \bar{\sigma}_i \tag{A.3}$$

$$L_i^h(\sigma) = \frac{\left[A_i(\sigma)\alpha^h(\sigma)\right]^{\epsilon-1} L_i^h}{\int_{\bar{\sigma}_i}^1 \left(A_i(\sigma')\alpha^h(\sigma')\right)^{\epsilon-1} d\sigma'}, \qquad \text{for all } \sigma > \bar{\sigma}_i. \tag{A.4}$$

The task threshold is endogenously determined in equilibrium by a 'no-arbitrage' condition that equalizes the cost of producing this threshold task using low or high skill workers. In general, workers of the same skill group receive the same city-specific wage in equilibrium even though they perform different tasks because no worker would supply their labor to tasks paying lower wages. Then, the threshold task $\bar{\sigma}_i$ must be such that this task can be produced with the same costs using either high or low skill workers. This is equivalent to task $\bar{\sigma}_i$ having the same equilibrium supply either when produced only with high or low skill workers. This implies the following no arbitrage condition between high and low skill workers for city i

$$\frac{\alpha^l(\bar{\sigma}_i)^\epsilon L_i^l}{\int_0^{\bar{\sigma}_i} \left(A_i(\sigma')\alpha^l(\sigma')\right)^{\epsilon-1} d\sigma'} = \frac{\alpha^h(\bar{\sigma}_i)^\epsilon L_i^h}{\int_{\bar{\sigma}_i}^1 \left(A_i(\sigma')\alpha^h(\sigma')\right)^{\epsilon-1} d\sigma'}.$$

$$\Rightarrow \qquad \frac{L_i^h}{L_i^l} = \left(\frac{\alpha^h(\bar{\sigma}_i)}{\alpha^l(\bar{\sigma}_i)}\right)^{-\epsilon} \left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)}\right)^{\epsilon-1}, \tag{A.5}$$

where $\Omega_i^l(\bar{\sigma}_i) \equiv [\int_0^{\bar{\sigma}_i}(A_i(\sigma')\alpha^l(\sigma'))^{\epsilon-1}d\sigma']^{\frac{1}{\epsilon-1}}$ and $\Omega_i^h(\bar{\sigma}_i) \equiv [\int_{\bar{\sigma}_i}^1(A_i(\sigma')\alpha^h(\sigma'))^{\epsilon-1}d\sigma']^{\frac{1}{\epsilon-1}}$. The task threshold is determined by the skill-specific and the city-specific productivity for producing tasks as well as the equilibrium ratio for high to low skill workers in a city.

Local labor demand. Comparing two tasks performed by low and high skill workers ($\sigma < \bar{\sigma}_i < \sigma'$), we obtain $p_i(\sigma)^{\epsilon} A_i(\sigma) \alpha^l(\sigma) L_i^l(\sigma) = p_i(\sigma')^{\epsilon} A_i(\sigma') \alpha^h(\sigma') L_i^h(\sigma')$. Thus, we must have for the relative wage of high skill workers

$$\frac{w_i^h}{w_i^l} = \left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)} \right)^{\frac{\epsilon-1}{\epsilon}} \left(\frac{L_i^h}{L_i^l} \right)^{-\frac{1}{\epsilon}}. \tag{A.6}$$

Rearranging this equation gives the relative local labor demand, conditional on the task threshold

$$\frac{L_i^h}{L_i^l} = \left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)} \right)^{\epsilon-1} \left(\frac{w_i^h}{w_i^l} \right)^{-\epsilon}. \tag{A.7}$$

Since the final good is the numéraire, we have $\int_0^1 \ln p_i(\sigma)^{1-\epsilon} d\sigma = 1$. This equilibrium condition can be written as

$$\int_0^{\bar{\sigma}_i} \left[\frac{w_i^l}{A_i(\sigma)\alpha^l(\sigma)} \right]^{1-\epsilon} d\sigma + \int_{\bar{\sigma}_i}^1 \left[\frac{w_i^h}{A_i(\sigma)\alpha^h(\sigma)} \right]^{1-\epsilon} d\sigma = 1.$$

Using the equation for the relative wage, this equating can be solved for the wage level of low and high skill workers (local labor demand)

$$w_i^l = \Omega_i^l(\bar{\sigma}_i) \left[1 + \left(\left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)} \right)^{\frac{1}{\epsilon-1}} \frac{L_i^h}{L_i^l} \right)^{\frac{1}{\epsilon}} \right] \tag{A.8}$$

$$w_i^h = \Omega_i^h(\bar{\sigma}_i) \left[1 + \left(\left(\frac{\Omega_i^h(\bar{\sigma}_i)}{\Omega_i^l(\bar{\sigma}_i)} \right)^{\frac{1}{\epsilon-1}} \frac{L_i^h}{L_i^l} \right)^{-\frac{1}{\epsilon}} \right]. \tag{A.9}$$

A.2 Local Labor Supply

Preferences. Each worker has Cobb-Douglas preferences for the consumption good and housing, and maximizes utility subject to the budget constraint. The utility of worker ω, who is of skill type s and lives in city i, is

$$\max_{c_i^s, h_i^s} U_i^s(\omega) = \frac{1}{\beta^{\beta} (1-\beta)^{1-\beta}} c_i^{s\,1-\beta} h_i^{s\,\beta} \epsilon_i^s(\omega) \qquad \text{s.t.} \qquad c_i^s + r_i h_i^s = w_i^s. \tag{A.10}$$

Here, c_i^s is consumption, h_i^s denotes housing with price r_i in city i, and $\epsilon_i^s(\omega)$ is an idiosyncratic preference shock for living in city i. This preference shock is independent and identically distributed across workers within the same skill group and across cities.

FOCs:

$$\frac{1-\beta}{\beta^{\beta}\left(1-\beta\right)^{1-\beta}} c_i^{s-\beta} h_i^{s\beta} \epsilon_i^s(\omega) = \mu \qquad \text{and} \qquad \frac{\beta}{\beta^{\beta}\left(1-\beta\right)^{1-\beta}} c_i^{s1-\beta} h_i^{s\beta-1} \epsilon_i^s(\omega) = \mu r_i$$

$$\Rightarrow \qquad c_i^s = \left(1-\beta\right) w_i^s \qquad \text{and} \qquad h_i^s = \frac{\beta}{r_i} w_i^s$$

The worker's maximized utility function can be expressed as an indirect utility function for living in city i as follows

$$V_i^s(\omega) = V_i^s \epsilon_i^s(\omega) = \frac{w_i^s}{r_i^{\beta}} \epsilon_i^s(\omega), \tag{A.11}$$

where V_i^s is the systematic indirect utility component. We assume that the idiosyncratic preference shocks $\epsilon_i^s(\omega)$ are Fréchet distributed with scale parameter $\kappa > 1$.

Local labor supply. The cumulative distribution function and the probability density function of the Fréchet distribution are given by

$$F_i^s(\epsilon) = e^{-\epsilon^{-\kappa}} \qquad \text{and} \qquad f_i^s(\epsilon) = \kappa \epsilon^{-\kappa-1} e^{-\epsilon^{-\kappa}}.$$

Since the indirect utility for a worker is a linear transformation of the idiosyncratic preference shock, the indirect utility is again Fréchet distributed with

$$F_i^s(v) = e^{-\Psi_i v^{-\kappa}}, \qquad \Psi_i \equiv \left(w_i^s\right)^{\kappa} \left(r_i^{\beta}\right)^{-\kappa}.$$

From all cities, each worker chooses the city that offers the maximum utility. Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the probability that a worker with skill s chooses city i is

$$\begin{aligned}
\lambda_i^s &= Pr\left(v_i^s \geq \max_{i' \neq i}\right) = \int_0^{\infty} \prod_{i' \neq i} F_{i'}^s(v) dF_i^s(v) = \int_0^{\infty} \prod_{i' \neq i} F_{i'}^s(v) f_i^s(v) dv \\
&= \int_0^{\infty} \prod_{i' \neq i} e^{-\Psi_{i'} v^{-\kappa}} \kappa \Psi_i v^{-\kappa-1} e^{-\Psi_i v^{-\kappa}} dv = \int_0^{\infty} \prod_{i' \in \mathcal{I}} e^{-\Psi_{i'} v^{-\kappa}} \kappa \Psi_i v^{-\kappa-1} dv \\
&= \int_0^{\infty} \kappa \Psi_i v^{-\kappa-1} e^{-\sum_{i' \in \mathcal{I}} \Psi_{i'} v^{-\kappa}} dv = \left[\frac{\Psi_i}{\sum_{i' \in \mathcal{I}} \Psi_{i'}} e^{-\sum_{i' \in \mathcal{I}} \Psi_{i'} v^{-\kappa}} dv\right]_0^{\infty} \\
&= \frac{\Psi_i}{\sum_{i' \in \mathcal{I}} \Psi_{i'}} = \frac{\left(\frac{w_i^s}{r_i^{\beta}}\right)^{\kappa}}{\sum_{i' \in \mathcal{I}} \left(\frac{w_{i'}^s}{r_{i'}^{\beta}}\right)^{\kappa}}
\end{aligned}$$

Thus, the share of workers with skill s who choose city i (local labor supply to the city) is given by

$$\frac{L_i^s}{L^s} = \frac{\left(\frac{w_i^s}{r_i^\beta}\right)^\kappa}{\sum_{i' \in \mathcal{I}} \left(\frac{w_{i'}^s}{r_{i'}^\beta}\right)^\kappa} = \frac{V_i^{s\kappa}}{\sum_{i' \in \mathcal{I}} V_{i'}^{s\kappa}} = \frac{V_i^{s\kappa}}{V^{s\kappa}}, \tag{A.12}$$

where $V^s \equiv \left(\sum_{j \in \mathcal{I}} V_j^{s\kappa}\right)^{\frac{1}{\kappa}}$ is the expected utility for a worker with skill s across all cities.

B Data Appendix

B.1 Data Description

We use the Regional File of the Sample of Integrated Labor Market Biographies (SIAB-R) provided by the Institute for Employment Research (IAB) of the German Federal Employment Agency (Antoni et al., 2019; Dauth and Eppelsheimer, 2020). This panel dataset provides a 2% random sample of administrative social security records from 1975 to 2017. It is representative of 80% of the German workforce and covers employees subject to social security contributions and marginal part-time employment but excludes the self-employed, civil servants, and individuals in military service. The dataset contains employment biographies of 1,827,903 individuals with a total of 62,340,521 observations. Because of anonymization, 2.5% of the individuals in the original SIAB are not contained in the Regional File.

Yearly panel and geographical unit of observation. The data is in spell format where the unit of observation is any change in the employment status of an individual. We convert the dataset into a yearly panel with June 30 as the cut-off date as the on establishment variables from the Establishment-History-Panel (BHP) are only exact on 30 June each year. The workplace of an individual is observable at the level of district regions which is our geographical unit of analysis. District regions are constructed from the administrative districts in Germany by aggregation to at least 100,000 inhabitants per district region. This yields 328 district regions of which 185 are 'urban' and 143 are 'rural'. As some smaller towns are located within rural district regions, we keep all district regions in the sample.

Wages. Due to a cap on social security contributions, wages are right-censored at the upper earnings limit for statutory pension insurance. Therefore, we impute wages above the censoring threshold following the procedure in Dustmann et al. (2009) and Card et al. (2013). We run a series of tobit imputations for each year, East-West Germany, and three education groups, separately. In each tobit estimation, we predict censored wages employing controls for age, gender, experience, and occupation groups. Furthermore, we include the individuals' mean wage in other years, the fraction of top-censored wages in other years, and a dummy if the individual is observed only once in the sample. Instead of including firm mean log-wages as in Card et al. (2013), we include annual mean log-wages in the district region. Under the assumption that wages are log-normally distributed, we impute censored log wages as follows: $X\beta + \sigma\Phi^{-1}\left[k + u(1-k)\right]$, $k = \Phi\left[(c - X\beta)/\sigma\right]$ where $u \sim U[0,1]$, s is the censoring limit, and σ is the standard deviation of the residual. Additionally, we drop wages of zero because they can be attributed to employment interruption notifications. During these periods, the employment relationship continues to exist in legal terms but without payment. We inflate wages to 2015 prices using the German consumer price index.

Definition of variables. We follow previous work that uses education as a proxy for skill (e.g., Acemoglu and Autor, 2011). Then, we define two skill groups: high skill workers (university or applied university degree) and low skill workers (vocational training or no vocational training). For better illustration, we aggregate the 120 occupations included the data into broader groups following the classification of Böhm et al. (2022) which is in the spirit of Acemoglu and Autor (2011). The resulting variable comprises the following four categories: managers, professionals, and technicians; sales and office workers; production workers, operators and craftsmen; and workers in services and care occupations. Following Eberle and Schmucker (2019), we construct the variables work experience and firm tenure. Work experience counts the number of days an individual was employed up to the current time point excluding training periods. Firm tenure measures the number of days an individual was employed in an establishment including training periods. Both variables are expressed in years.

Sample restriction. We restrict the sample to the time span between 1980 and 2010 and to West Germany only. Until 2010, the SIAB Regional File includes 120 occupations which are consistently defined over time. Because the classification of occupations changes drastically after 2010, we cannot use subsequent years. For historical reasons, there are considerable differences in wages between East and West Germany and complete data for East Germany becomes available only after 1993. Further, we constrain the sample to men in full-time employment liable to social security between the age of 18 and 62. Because only the wage but not the underlying hours worked are observed for part-time workers, it is not possible to control for changes in part-time workers' wages due to changes in the underlying hours worked. Finally, we drop all observations for which the information on education, occupation or workplace location is missing. The final sample counts 6,539,677 observations.

B.2 Further Descriptive Statistics

Table B1: Descriptive statistics on employment shares by occupation groups, 1980-2010

	Level (%)				Change (%)
	1980	1990	2000	2010	1980-2010
Panel A: All cities					
Managers-Professionals-Technicians	16.91	19.25	21.93	24.24	+7.33
Sales-Office	15.26	15.38	17.07	18.25	+2.99
Production-Operators-Craftsmen	63.44	60.33	55.21	51.26	-12.18
Service-Care	4.38	5.03	5.78	6.25	+1.87
Panel B: Top 33% of cities					
Managers-Professionals-Technicians	20.36	23.15	26.70	29.24	+8.88
Sales-Office	17.17	17.59	19.80	21.63	+4.46
Production-Operators-Craftsmen	58.00	54.06	47.29	42.36	-15.64
Service-Care	4.48	5.20	6.21	6.76	+2.28
Panel C: Middle 33% of cities					
Managers-Professionals-Technicians	13.24	15.55	17.65	20.10	+6.86
Sales-Office	12.95	12.76	14.18	14.82	+1.87
Production-Operators-Craftsmen	69.81	67.19	63.27	59.66	-10.15
Service-Care	3.99	4.49	4.90	5.42	+1.43
Panel D: Bottom 33% of cities					
Managers-Professionals-Technicians	10.85	12.60	14.80	16.83	+5.98
Sales-Office	12.31	12.38	13.61	14.10	+1.79
Production-Operators-Craftsmen	72.25	69.73	65.73	63.05	-9.20
Service-Care	4.60	5.29	5.85	6.02	+1.42

Notes: Occupation groups are defined according to the classification of Böhm et al. (2022). In panels B to D, the sample is split into three groups of cities by population density in 2011 such that each subsample includes an equal number of cities. Panel B includes the one-third of cities with the highest population density, panel C the ones in the middle, and panel D the cities with the lowest population density. See section B.1 for further details on the data. All numbers are in percent.