Why has technological change not closed the gender wage gap?

Estefanía Galván*

Abstract

The impact of technological change and the extent to which labor markets are increasingly polarized are actively debated in current research. However, we know little about the effect of these structural changes on the dynamics of the gender wage differentials. Using administrative panel data for Germany, this paper investigates how the changes in wages caused by technological progress have affected the gender wage gap. I study this question by estimating changes over time in occupation wage premiums and decomposing its effects into gender differences in sorting and wages changes across occupations associated with advancing technology. My results show that the effect of gender segregation has mostly benefited female workers, contributing to narrow the gender gap. This is mostly explained by the fact that men are over-represented in manual routine occupations, which leave them more exposed to automation of work, but also because women increased their employment in better-pay cognitive non-routine occupations. However, this effect has been offset by gender differences in wage changes within occupations. I find that within the cognitive groups the wage gains over time for male workers have been larger than those for females. This explains why, although women have been less exposed to the automation of work and increased their employment in non-routine high-pay occupations, technological change did not lead to greater reductions in the gender wage gap.

Keywords: gender gaps, occupations, wage premia, technological change
JEL Classification: D13, J16, J31

*IECON, Universidad de la República. Gonzalo Ramirez 1926 - C.P. 11200. Montevideo, Uruguay. Email: estefania.galvan@fcea.edu.uy. This paper is part of my PhD thesis conducted at Aix-Marseille University, AMSE, France. I am grateful to Cecilia García-Peñalosa, Christian Schluter, Christina Gathmann, Andrea Ichino and Libertad González for substantial comments on earlier drafts. I acknowledge participants at AMSE, IECON and dECON seminars, LACEA-LAMES Annual meeting, LISER-IAB Conference, 9th ECINEQ Conference, LAGV Conference and the 34th Annual Conference of ESPE, for useful comments and suggestions. This work was supported by the *Agencia Nacional de Investigación e Innovación* (ANII - Uruguay) Grant POS_EXT_2015_1_125059; and the French National Research Agency Grant ANR-17-EURE-0020. This study uses the weakly anonymous Sample of Integrated Labour Market Biographies (SIAB 7514). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access.

1 Introduction

There is an intense debate worldwide on the impacts of technological progress on labor markets. One of the most documented consequences of these structural changes is the relocation of employment from routine work to non-routine tasks. To the extent that these routines are based on well-defined rules and procedures, they can be carried out by a computer running a program (automation), reducing demand in medium-grade occupations, a phenomenon that has been called routine biased technological change (RBTC).[1] This has led to a growing literature investigating the effect of technological change on inequality and the extent to which labor markets are increasingly polarized (Autor et al., 2003; Goos et al., 2009; Autor et al., 2008). However, we know little about how technological progress has affected the dynamics of the gender wage gap.

Although these structural transformations are in principle "gender-neutral", they might have had relevant implications for the evolution of gender wage disparities if men and women have different labor supply patterns across occupations that were differently affected by technological change. With the advancement of automation, the occupations that lost the most in terms of employment and wage returns are traditionally male industrial-goods-producing occupations, which have a high content of routine tasks. In this context, by looking at the employment structure, a recent literature suggests that female workers have relatively benefited from technological change (Cortes et al., 2018; Ngai and Petrongolo, 2017; Borghans et al., 2014; Black and Spitz-Oener, 2010). Despite this, the convergence in the gender wage gap stagnated in the last decades (Blau and Kahn, 2017; Olivetti and Petrongolo, 2014), putting in question the extent to which the impact of technological change on occupational wages has been favorable to female workers.

Using administrative data for Germany, this paper investigates how changes in wage dynamics and in the employment structure across occupations associated with technological progress have affected the gender wage differences over time. The German labor market presents a particularly interesting case. First, Germany has one of the highest and most persistent gender wage gaps among developed countries (Olivetti and Petrongolo, 2017), and it is one of the high-income countries that experienced the largest increases in female labor force participation in recent decades.[2] Moreover, the changes in the composition of female employment show a shift towards high-wage occupations, in line with a female labor force that is relatively more educated than before, which should have allowed women to benefit from the increased skill premiums. Secondly, having a large industrial sector, the effect of technological change on the employment structure has been remarkable (Dustmann et al., 2009; Spitz-Oener, 2006). In fact, Dustmann et al. (2009) find that technological change is an important driving force behind the widening of the wage distribution observed in Germany since the 1990s.[3] With wages in traditional male industrial

[1]While the literature has provided different explanations for polarization patterns, Goos et al. (2014) find that for European countries, RBTC is much more important than offshoring.

[2]According to OECD (2017) it increased from 38.6% in 1970 to 55.6% in 2016.

[3]They show that technology does not simply increase the demand for skilled labor relative to that of unskilled labor, but instead asymmetrically affects the bottom and the top of the wage distribution. While during the 1980s the increase in wage inequality was concentrated at the top of the distribution, in the 1990s it occurred at the bottom as well. They find that polarization also affected female workers, however, they restrict the analysis

occupations falling we would expect to see reductions in the gender wage gaps in recent decades, yet these gaps have stagnated since the 1990s.

Therefore, the main substantive contribution of this paper is to investigate why is it that the large wage gains that are observed in the occupations in which women are increasingly employed in detriment of male dominated occupations have not led to a further reduction in the gender wage gap. To shed light on this question I study the effect of technological change on the wage gap dynamics by analyzing the role played by occupational sorting of male and female workers and the differential changes in wage trajectories across occupations that were differently affected by RBTC. More specifically, I estimate changes in occupation-specific wage premiums, which are defined as the component of the worker's potential wage that is common to all workers in each occupation group in a given year, after accounting for the effect of occupation-specific returns to the individual's skills. Then I investigate how these differential changes in wage premiums for male and female workers affect the gender wage gap given their sorting across occupation groups.

To estimate time-varying occupation wage premiums I follow the panel data approach developed by Cortes (2016). The advantage of this empirical method is that it allows to control for the self-selection of workers into occupations based on observable and unobservable individual characteristics, therefore addressing concerns related to the effect of endogenous selection into occupations. My main dataset is the German *Sample of Integrated Labour Market Biographies (SIAB)*, which is particularly suitable for assessing the effect of technological change over time on wage trajectories, as it allows us to track individual labor biographies over the entire career. I combine this database with survey information on task content of jobs to characterize the occupations in which individuals are employed. I consider a classification of five broad occupation groups, which is based on the nature of the tasks mainly involved, distinguishing between tasks which can be automated, and tasks that require analyzing or interacting with others, and where technology often complements work. These occupation groups are: analytical non-routine (e.g. engineers and researchers), interactive non-routine (e.g. managers and teachers), cognitive routine (e.g. secretaries and clerical workers), manual routine (e.g. assemblers), and manual non-routine (e.g. cleaners and repairing service).

My results show that the effect of the gender segregation across occupations has mostly benefited female workers, contributing to narrow the gender wage gap. This is explained both because men are over represented in manual routine occupations (mainly industrial blue-collar occupations), which leave them more exposed to automation of work, and because women have moved out of manual routine work, moving to better paid cognitive non-routine occupations. However, the effect of these structural changes on the convergence of the gender wage gap was attenuated by the fact that wage growth for female workers within those occupations that were most favored by technological change was lower than that of men. By estimating gender-specific changes in wage premiums across occupations, I find that the premiums for males grew more rapidly than those for female workers within cognitive non-routine occupations and that

of the rising wage inequality to male workers.

these within-occupation differences in wage premiums growth largely dominate the compensating effect of gender differences in sorting across occupations. From 1992 to 2010 the gender gap in occupation premiums rose on average by 14 log points. This result implies that absent these gender differences in the changes of occupation premiums over time, the gender wage gap would have declined by 35% of the 1992 wage gap, rather than the 10% that we observe. Moreover, I find that the gender divergence in wage premiums over time is more important for the most recent cohorts, which suggests that these effects are not likely to be reverted in near future.

This paper contributes to the literature that aims to understand the causes of the persistent gender inequalities in the labor market. The roots of the stagnation in the narrowing of the gender wage gaps are actively debated in current labor market research. After prolonged gender convergence in education and experience, the explanatory relevance of human capital variables has decreased. On the contrary, despite the occupational upgrading of women relative to men, occupational structure remains an important factor in explaining the gender wage gaps (Goldin, 2014; Goldin et al., 2006). For example, Blau and Kahn (2017) based on an Oaxaca-Blinder decomposition for the US, find that occupation and industry are now the single largest measured variables accounting for the gender pay gap: they accounted for 20% of the gap in 1980 and for 51% of the (smaller) 2010 gap. The picture is qualitatively similar in the UK, where Petrongolo and Ronchi (2020) find that industry and occupation controls jointly explain 44% of the 2017 wage gap. This suggests that structural changes that alter the employment structure, the skills required, or the wages paid across occupations might be relevant factors to understand the evolution of the wage gap.

In this context, this paper is related to a strand of literature focusing on the effect of structural changes across occupations on the gender wage gaps. This literature has focused mainly on the employment structure side, that is, analyzing to what extent women's jobs are more or less subject to automation compared to those of men, and it suggests that female workers have relatively benefited from technological change. For example, Ngai and Petrongolo (2017) identify the reallocation of labor from goods to service industries as a primary driver of the rise of female relative hours of work and wage gains, by creating creating jobs that are less physically demanding and more intense in interpersonal skills. Moreover, focusing on the task composition of occupations, some studies show that this was favored by women's comparative advantages in the use of communication and interpersonal skills, that cannot be easily automated (Borghans et al., 2014; Cortes et al., 2018), in line with a recent experimental literature highlighting some gender differences in social attitudes such as altruism, fairness and caring behavior (see Azmat and Petrongolo (2014) for a review). As these social skills are increasingly valued in the labor market, women experienced large rises on the probability of working in a cognitive high-wage occupation.

In particular, my paper is closely related to Black and Spitz-Oener (2010), which focus on low and medium educated workers in West Germany and analyze changes in the task composition of the work of men and women between 1979 and 1999. They show that women have witnessed relative increases in analytic and interactive tasks within occupation and industry cells, and a strong decline in routine tasks, which they interpret as a positive effect of technological change on female workers relative to males. In line with them, I find that medium-skill male workers were those most affected by technological change, as my estimates show decreasing wage premiums over all the period in manual routine occupations (in which men are over-represented). However, unlike previous studies, this paper investigates the effects of technological advances on the gender wage gap by looking not only at the employment structure side, but also to the effect on wage changes across occupations. By investigating male and female wage trajectories across occupation groups, I find that the wages for male workers grow more rapidly than those of female workers within non-routine analytical and interactive occupations, which explains why despite gaining participation in these increasingly valued tasks, technological change has not led to further reductions in the gender wage gap. Moreover, the effect is strongest for most recent cohorts, which suggest that women still face some constrains that did not allow them to benefit from the increased overall wage returns in the upper part of the skill distribution.

These results are in line with recent work by Cortes et al. (2020) who, using survey data for Portugal and the US find that while women have been less exposed to the automation of work, at times they have reallocated to jobs with lower wage levels, and therefore technological change has not always led to declining wage gaps. Unlike them, this paper uses individual panel data and focuses on the wage changes over time across occupations associated with advancing technology, contributing to close a gap in the literature, which has generally investigated the effect of RBTC in the labor market at an aggregate level, relying on cross-sectional comparisons of occupational composition of employment across decades.[4]

The remainder of the paper is structured as follows. Section 2 introduces the empirical approach. Section 3 presents the data and discusses some descriptive evidence on labor market trends in wages and gender composition of employment across occupations. In Section 4 I present the main results on the gender differences in the changes in occupation premiums and the contribution to explain the wage gaps trajectories, while in Section 5 I investigate changes over the life-cycle and potential cohorts effects. Section 6 conducts a series of robustness checks on the results. Next, Section 7 discusses further evidence that provides some insights on why wage gains for men grow more rapidly than those of females within the same occupation groups. Lastly, Section 8 concludes.

2 Empirical approach

2.1 Estimation of gender-specific time-varying occupational wage premiums

To estimate the effect of RBTC on wage changes, I follow the empirical approach of Cortes (2016), who developed a method for the unbiased and consistent estimation of changes over time in occupational wage premia after controlling for selection into occupations based on observable and unobservable individual characteristics.

[4]Notably exceptions to this are Cortes (2016) for the US and Bachmann et al. (2019) for Germany. Both focus on male workers who are initially in routine jobs and analyze their wage trajectories and mobility patterns both as stayers and switching-out of routine occupations.

The underlying theoretical model assumes that there is a continuum of workers, who differ in terms of their skill levels. There is perfect information and workers sort endogenously into one of the broad occupation groups. In this paper I consider five broad occupation groups: analytical non-routine (ANR), interactive non-routine (INR), cognitive routine (CR), manual routine (MR) and manual non-routine (MNR).[5] Occupational sorting is driven by comparative advantage (as in Gibbons et al., 2005). Workers of higher skill levels are more productive at all tasks, but particularly so at more complex tasks. Potential wages for each worker in each occupation are the product of the competitively determined wage per efficiency unit in that occupation and the number of efficiency unit supplied by the worker. In equilibrium, there would be endogenously determined skill thresholds that determine the optimal selection into occupations for each worker. That is, according to his individual skill level each worker will select into one of the five occupation groups so that the least skilled workers find it optimal to select into the manual non-routine occupations, while the most skilled workers into analytical non-routine occupations. The cutoffs are determined in equilibrium so that the marginal workers have no incentives to relocate between task-occupational groups. The demand for cognitive non-routine occupations is relatively low, making optimal only for the most skilled workers to select into these occupations (where they are much more productive), while the least skilled workers are attracted to the manual non-routine occupations (as their extra productivity in the other tasks is relatively small). The equilibrium distribution implies that wages will be on average lowest among manual non-routine workers and highest among cognitive non-routine (which is consistent with the data on Germany).

From this model, the potential wage for an individual of skill level z_i in occupation j (where $j= \{ANR, INR, CR, MR, MNR\}$), consist of an occupation wage premium, which is common to everyone in the occupation, and on the individual's occupation specific productivity. Assuming that productivity is log-linear in skills, we have the following equation for the potential log wage for individual i in year t:

$$w_{ijt} = \theta_{jt} + z_i a_j \tag{1}$$

where θ_{jt} is the occupational wage premium in occupation j at time t, z_i is the skill level of individual i and a_j may be interpreted as an occupation-specific return to skills. These returns to skills (a_j) vary across occupations, and following the assumptions of comparative advantage from the model, we assume that they are highest in cognitive non-routine occupations and lowest in manual non-routine. Then, there will exist critical values of the skill levels (z_i) that would lead workers of different abilities to self-select into the occupations where each of them can have the highest return, determining an efficient assignment of workers to occupations.

In the equation to be estimated empirically, the observed wage will depend on the occupation in which the individual is employed:

[5] The details on the classification of detailed occupations on these broad groups, which are based on the nature of the tasks involved, are given in Section 3.

$$w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}\gamma_{ij} + u_{it} \tag{2}$$

where D_{ijt} is an occupation selection indicator that equals one if individual i selects into the occupation j at time t, and $\gamma_{ij} = z_i a_j$, which can be interpreted as an occupation-spell fixed effect for each individual. Initially I will assume that individual skills (z_i), as well as the return to skills (a_j) are time-invariant, that is, fixed over their lifetime. Then, γ_{ij} varies across occupation spells but it stays constant whenever an individual stays in the same occupational category.

Because individual's skills and the occupation-specific returns to skills are not varying over time, occupational mobility will be driven exclusively by changes over time in the occupation premiums. Then, this model assumes comparative advantage in the sorting into occupations, but mobility is exogenous (driven by RBTC). Workers who do not find it optimal to switch occupation experience a wage change equal to the change in the wages per efficiency unit in their optimal occupation. The identifying assumption is that selection into occupations only depends on occupation fixed-effects and individual worker's ability.[6] That is, conditional on these two elements, selection into each occupation group is random: $E(u_{it}|D_{ijt}, z_i, \theta_{jt}) = 0$.[7] In practice the occupation-spell fixed effect will capture not only individual ability or skills, but also wage effects of all time-invariant characteristics of the individual that affect wages within the occupation spell.

These assumptions rule out dynamic effects, such as workers learning about their ability over time or that individuals move from learning occupations to earning occupations. In the latter case, individuals first accumulate experience/human capital but earn lower wages; they then switch over their career to high-wage occupations to reap the benefits of their investment. In the robustness section (Section 6) I relax some of the assumptions, introducing changing returns over time to education (that is an observable component of the ability) and occupation-specific tenure profiles, which allows for heterogeneous returns to occupational tenure. That is, the possibility that the tenure profile is steeper in non-routine analytical and interactive occupations than in routine ones.

I am interested in allowing occupation wage premia to differ by gender, that is, estimating gender-specific wage premiums for each occupation group. With this objective, I incorporate an interaction between the year-occupation fixed effects and a dummy that takes value one for female workers (fem_i). The regression being estimated is therefore:

$$w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}\beta_{jt}\, fem_i + \sum_j D_{ijt}\gamma_{ij} + X_{it}\delta + u_{it} \tag{3}$$

[6] In Section 6 I allow for occupation-specific tenure profiles. Then, the individual's occupational choice will also depend on his/her tenure in the current occupation.

[7] Cortes (2016) recognizes that in practice occupational mobility is affected by search frictions that restrict workers from immediately selecting into the most convenient occupation in each year. In this sense, the identifying assumption would be that conditional on occupational fixed-effects and individual worker's unobserved ability selection into occupation is only driven by a search friction that is orthogonal to ability or other wage determinant. If this assumption holds, the coefficients in equation (2) are consistently estimated.

where I have also added a vector of additional variables, X_{it}, which include year fixed effects and a set of controls: dummies for region of work at the federal state, experience and a dummy that takes value one if the individual is a German national.[8]

I define $\theta^M_{jt} \equiv \theta_{jt}$ and $\theta^F_{jt} \equiv \theta_{jt} + \beta_{jt}$ as the occupation-year fixed effects for male and female workers respectively. In the estimation standard errors are clustered at the individual level. This equation can be consistently estimated using fixed effects at the occupation-spell level for each individual, that is, using a fixed effect for each individual in each occupation in which he/she is observed in. θ^M_{jt} and θ^F_{jt} are estimated through interactions of occupations and year dummies. The omitted category is the manual non-routine for male workers. The inclusion of year dummies captures changes over time that affect workers in all occupations. Because of the inclusion of the occupation-spell fixed effects, the occupation-time fixed effects are identified only from variation over time within occupation spells. Therefore the estimates $\hat{\theta}^g_{ANRt}$, $\hat{\theta}^g_{INRt}$, $\hat{\theta}^g_{CRt}$ and $\hat{\theta}^g_{MRt}$ with $g = \{M, F\}$ should be interpreted as a double difference: they identify changes over time in the occupational wage premium relative to the base year and relative to the changes experienced by the base category (manual non-routine male workers).

2.2 The role of changes in occupation premiums for explaining the gender wage gaps

Using the estimated changes in the occupation-year fixed effects coming from equation (3), I analyze the role of occupation premiums for the gender wage gap. Denoting $E\left[\hat{\theta}^M_{jt}|male\right]$ as the average change in the wage premium received by men in occupation j, that is, weighting the occupation premiums by the male distribution of employment across occupations, and $E\left[\hat{\theta}^F_{jt}|female\right]$ as the average change in the wage premium received by women in occupation j, the difference between male and female average change in wage premiums is given by the following equation:

$$Gap\,\bar{\hat{\theta}}_t = E\left[\hat{\theta}^M_{jt}|male\right] - E\left[\hat{\theta}^F_{jt}|female\right] = \sum_j s^M_{jt}\hat{\theta}^M_{jt} - \sum_j s^F_{jt}\hat{\theta}^F_{jt} \tag{4}$$

where s^M_{jt} and s^F_{jt} are the proportion of male and female workers employed in occupation j in year t.[9]

Two complementary channels determine how changes in occupation-specific wage premiums affect the gender wage gap: a composition or sorting across occupations channel and a component that measures gender differences within occupations. The first channel takes place if women are less likely to be employed at higher-wage occupations, while the within occupation differences arise if women obtain a smaller occupation premium than men for the same occupation group.

To analyze the relative importance of these two explanatory channels, I follow the approach of Card et al. (2016), who perform a Oaxaca-Blinder decomposition (Oaxaca, 1973; Blinder,

[8]It is assumed that these control variables are orthogonal to the measurement error u_{it} and that their effects are not occupation-specific.

[9]The gender gap in the estimated occupation premiums ($Gap\,\hat{\theta}_{jt}$) is invariant to which occupation category is used as reference.

1973):[10]

$$Gap\,\bar{\hat{\theta}}_t = E\left[\hat{\theta}^M_{jt} - \hat{\theta}^F_{jt}|male\right] + E\left[\hat{\theta}^F_{jt}|male\right] - E\left[\hat{\theta}^F_{jt}|female\right] \tag{5}$$

$$= E\left[\hat{\theta}^M_{jt} - \hat{\theta}^F_{jt}|female\right] + E\left[\hat{\theta}^M_{jt}|male\right] - E\left[\hat{\theta}^M_{jt}|female\right] \tag{6}$$

The first term in equations (5) and (6) is the average within occupation effect, calculated by comparing the changes in the occupation premiums for male and female workers across the distributions of occupations held by men (5) or by women (6). It measures by how much the gender wage gap would change if women received the same change in the occupation premium as men. The second term is the sorting across occupations component, that measures by how much the gender wage gap would change if women had the same distribution across occupations than men, weighted by the change in the female (5) or male (6) occupation effects.

Summing up my empirical approach, I first estimate overall changes in the wage premiums across occupations and analyze its implications on the average change in the gender wage gap; secondly, I estimate gender-specific wage premiums and decompose the average gender gap in occupation premiums into two explanatory channels: sorting-across-occupations and within-occupation differences. I perform this analysis for the whole sample of male and female workers from 1975 to 2010, as well as in two sub-periods, and on the other hand, I estimate changes over time in the occupation premiums separately for four cohorts of workers.

3 Data and descriptive overview

3.1 Administrative data on labor market biographies

The data basis of this paper is the weakly anonymous *Sample of Integrated Labour Market Biographies* (SIAB) 1975 - 2014.[11] This large administrative dataset for Germany is a two percent random sample drawn from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). The data on labor market biographies consists of mandatory notifications made by employers to social security agencies, and therefore they contain information of jobs subject to social security contributions. This means that self-employed individuals, civil servants, and family workers are not included. For this reason, this data set represents approximately 80 percent of the German workforce.

In this study I focus on the period 1975-2010.[12] The data set provides information about the characteristics of the jobs held by workers and about some of their personal characteristics.

[10] This decomposition is inspired on Card et al. (2016), who proposed a similar approach to decompose the differences in male and female firm premiums, obtained from and AKM model (Abowd et al., 1999), into sorting across firms and bargaining power components.

[11] See Antoni et al. (2016) for details on data documentation.

[12] Even though the data contains information up to 2014, the reason I consider the period 1975-2010 is that since 2011 a new occupation code was introduced, which imply some changes in the way the data on labor histories

In particular, it contains information on gross daily wage, the number of days worked in a given year, the type of contract (whether full-time or part-time), the region of work (federal and district levels), the educational level, the gender and the year of birth, among others. For each worker in the dataset it contains an establishment identifier which can be used to match with information coming from the establishment side of the dataset. The SIAB tracks individual over time, making it possible to document all transitions between employment and non-employment, as well as direct job-to-job transitions using the establishment identification number.

Administrative data have several advantages, like a large number of observations, no non-response burden and no problems with interviewer effects or survey bias.[13] While the data is virtually free from measurement errors, there are two major shortcomings. First, these data are right-censored at the contribution assessment ceiling for the pension insurance. In Germany, employees contribute a share of their gross wage to the mandatory pension system up to a wage ceiling. As a result, information about wages in the SIAB is top-coded or right-censored at the upper limit of the social security system. In order to better approximate the true distribution of top earnings, I impute the wages of the individuals affected by top coding by using the heteroscedastic single imputation approach described by Büttner and Rässler (2008).[14] The second disadvantage is that it does not contain precise information on the number of hours worked. We only know if an individual is working part-time or full-time, defined as working at least 30 hours per week. Therefore, I restrict my main analysis to full-time workers so that wages are comparable. As the wage variable I use the daily wage, transformed into real daily wage at prices of 2010 by using the Consumer Price Index.[15] Excluding part-time workers from the analysis might be a concern since part-time work is quite frequent in Germany, notably for female workers. Therefore, in Appendix B I conduct a robustness check extending the analysis to include part-time workers. For that extension I used data from the German Socio Economic Panel (GSOEP) which contains information on hours worked and allows me to obtain an approximate hourly wage respectively for part-time and full-time workers in the SIAB data.

The data set contains information on school leaving qualification and vocational training.[16] Using these variables I create a variable for the skill level. The low skill level comprises individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications. The medium skill level includes those with a lower secondary,

was reported. The new occupation code in 2011 led to a number of problems. For example, during the transition period granted to employers in the social security notification procedure, there was a temporary increase in the number of missing details, which had to be solved by imputing the missing values (see See Antoni et al. (2016)). One of the most relevant changes is the switch from the Classification of Occupations 1988 (*Klassifikation der Berufe* 1988 - KldB 1988), to the KldB 2010. Although employment notifications after 2011 (KldB2010) are recoded to the KldB1988 by transferring the key area, it results in substantial inaccuracies. Given the relevance that the occupational classification has for this study, I consider the period until 2010.

[13] Compared to popular survey data sets like the SOEP, the main advantages of the the SIAB are its large size, the long period it covers, the almost complete absence of panel mortality or attrition and the reliability of the core variables like date and length of spells and wages (Fitzenberger et al., 2005).

[14] One of the advantages of the method they propose is that it does not presume homoscedasticity of the residuals. The estimation of imputed wages is conducted separately for each gender and year (more details are presented in the Appendix).

[15] German Federal Statistical Office. https://www.destatis.de

[16] See Appendix C for more details on education variables and the imputation of missing.

intermediate secondary or upper secondary school leaving certificate and a vocational qualification. The high skilled encompasses all employees who have a degree from a university of applied sciences (*Fachhochschule*), technical college degree or university degree.

My sample is composed of male and female workers between 25 and 55 years old in West Germany for the period 1975-2010, excluding apprentices and marginal part-time employees.[17] For the analysis, I create a dataset that contains one observation per year per worker. In the cases where there are overlapping observations for the same period and worker, I keep the job that has the highest wage. When there is more than one observation per worker in the same year (different periods) in different establishments, I keep the observation that correspond to the longest period, that is, the one with highest number of days worked in that year.

Table 1 records summary statistics for the sample of analysis. I have 7,102,028 person-year observations for men and 4,761,753 for women, which correspond to 565,230 individuals for men and 446,612 female workers. There is almost no difference in mean age between men and women (39 years old), and in experience (almost 12 years on average).[18] Most of the individuals in the sample are German nationals (89% of men and 92% for women). 33% of the women are part-time workers compared to only 2% of men. Most of the individuals (around 75%) have medium level skill education (qualified), that is, have at least a vocational qualification. 15% of the men and 11% of the women are highly qualified, that is, have a university or technical university degree.

3.2 Occupation groups based on task content of occupations

The classification of occupations that I use in this paper follows the classification developed by Spitz-Oener (2006), which is based on the task content of occupations. The term task refers to activities that individuals have to perform in their work. The task based approach (TBA) provides a conceptual basis for the comparison of different occupations, allowing to reduce the complexity of theoretically distinguishable dimensions of tasks to a few key dimensions that are most important for relevant research questions. Focusing on occupations groups classified based on the TBA has become standard in the literature, which is partly justified by the fact that many skills are occupation-specific (see for example Gathmann and Schönberg (2010)).

In Germany, the operationalisation first proposed by Spitz-Oener (2006) of the task-based framework (introduced by Autor et al. (2003)) is usually applied. Instead of expert evaluations, this approach is based on survey-based information on tasks that individuals do in their jobs. She proposes a five-task dimension classification: 1. Analytical non-routine (e.g. researching and analyzing), 2. Interactive non-routine (e.g. managing and teaching), 3. Cognitive routine

[17] Information on East Germany is available since 1992 in the SIAB. However, as I include the pre-unification period in the analysis, I restrict the sample to West Germany to have a consistent sample. Marginal part-time employees are included in the data from 1999 onward. 13.8% of the observations are below the marginal part-time income threshold. I exclude them from my sample (for this I used the values provided by the FDZ for the marginal part-time income threshold in each year and exclude the observations where the wage is above these values). I follow Dustmann et al. (2009) in excluding spells of workers in apprentice-ship training. Given that the ultimate focus is on the prime-aged population (25-55 years old), information loss due to dropping apprentices is negligible.

[18] Working experience is measured as the difference between the date of the observation and the date in which the individual has his /her first register as worker in the data (first day of work), divided by 365.24 to express it in years.

(e.g. calculating and bookkeeping), 4. Manual routine (e.g. operating machines), and 5. Manual non-routine (e.g. serving and repairing).

The distinction between routine and non-routine tasks refers to the fact that routines are based on well-defined rules and procedures, and therefore could be potentially carried out by a computer that executes a program (automation). Analytical refers to the necessity to think and analyze during work, whereas interactive denotes the need to communicate with others by oral or written means, ranging from dealing with co-workers or clients to complex interactive activities such as counseling, educating or teaching (Spitz-Oener, 2006).

To operationalize the task categories, I use data on tasks coming from the Qualification and Career Survey, which is a survey of employees carried out by the German Federal Institute for Vocational Training (BIBB). It includes six cross sections launched in 1979, 1985/86, 1991/92, 1998/99, 2006 and 2012, each covering a representative sample of about 30,000 workers (men and women). From these data I use worker self-reports on the tasks involved in their present job from a given list of activities. These tasks have been changing across waves. In order to create a task intensity measure that is consistent over time, I followed previous work and merged some of the activities in order to deal with the changing definitions of the variables and to maintain a total number of activities which is similar in each survey. I arrived to 17 longitudinally consistent tasks, and classified them into the five dimensions proposed by Spitz-Oener (2006) (see Table C.1 in the Appendix).

To construct a single measure of the different tasks performed by individuals it is necessary to calculate an index of task intensity. In previous work, two different task intensity measures were developed using these data. Spitz-Oener (2006) proposed an index of "task intensity", which determines the degree to which a single task dimension is necessary, to perform a specific occupational activity when compared to another occupational activity. Antonczyk et al. (2009) developed an index of "task composition", which specifies the shares of the different tasks in an individual's occupational activity. These indices are sensitive to the number of task variables included. Rohrbach-Schmidt and Tiemann (2013) discuss the sensitivity of using different approaches to aggregate the multiple questions related with tasks in the BIBB data across waves into time-consist variables and on how to classify them into the five broad groups.

I calculate both indices and take means of each category at the occupational level (3 digits of Kldb 1988) in order to be able to classify occupations in the SIAB data. The final classification is based on the highest value of each index at the occupational level (more details are presented in Section C of the Appendix). Tables C.5 to C.8 in the Appendix contain the detailed mapping of occupations into the task-based occupation categories. Contrary to Black and Spitz-Oener (2010) that use BIBB data to update task intensities and obtain time varying task-intensity on occupations and industry cells, in this paper I consider a task-based classification of occupations that is fixed over the period of analysis, as the main interest is to characterize the nature of occupations, and analyze gender differences in changes on wage premia and employment across occupations.

Figure 2 shows the evolution of the mean logarithms of real daily wage for full time workers by occupational group. As expected, we can see that according to this classification, analytical non-routine are on average high wage occupations, followed by interactive non-routine. Cognitive routine occupations are middle-wage, and manual occupations are low-pay (manual non-routine is the lowest-pay occupation for women). For each of these occupational groups the mean of the wage levels is always lower for female workers than for males. Along the period 1975-2010, the mean in log real wage in all cognitive occupations show a growing tendency for both men and women. For manual occupations, on the contrary, it grows until 1990 and then shows a slight decrease.

3.3 Trends in gender wage gaps and occupations

Before investigating the changes in occupation premiums over time, this subsection discusses the major trends in the evolution of employment and wages across occupations. In Table 1 we can observe a markedly segregation by gender across occupational groups. In this period, on average half of the men in the sample are employed in manual routine occupations (52%) compared to only 14% of women. Women are mostly employed in cognitive routine occupations (39%) and manual non-routine (31%). Considering the high skilled cognitive occupations, men are more represented in analytical (14% vs. 4% for women), while women are more represented in interactive (12% vs 8%).[19]

Figure 1 shows the variations in employment shares for each group over the period 1975-2010, and in two sub-periods to see how these patterns differ over time. The cumulative long-run change shows a sharp decline, of about 15% in the share of manual routine occupations for both men and women. For men, the decline in the share of manual routine was compensated by an increase in the share of workers in the other categories, and in particular in both extremes of non-routine occupations: analytical non-routine and manual non-routine. For women, the decline took place not only in routine manual occupations, but also in manual non-routine, and in the last half of the period (1992-2010) also the share of cognitive routine experienced a reduction in its relevance. These declines were compensated by increases in the share of female employment in non-routine occupations, and in particular, by a strong increase in the share of employment in interactive non-routine occupations (more than 10%). In fact, the interactive non-routine group and the analytical non-routine show a growing pattern for women along the two sub-periods considered. That is, women switch out of manual routine jobs moving disproportionately to cognitive non-routine occupations.

These patterns are in line with the hypothesis of routine biased technological change, which predicts a shift from occupations that involve routine tasks towards those that involve mostly analytic and interactive non-routine tasks. It is also consistent with previous evidence for Germany, by Black and Spitz-Oener (2010), who find that women witnessed relative increases in

[19] If we consider all the employees in these two occupation groups, the percentage of females in analytical occupations is only 16.5% while it is 49.6% in interactive occupations (Table C.4 in the Appendix). That is, male workers are overrepresented in both high skilled cognitive occupational groups, and especially in analytical occupations. The percentage of female workers, on the other hand, is higher than that of men in cognitive routine (67.6%) and manual non-routine occupations (60.3%).

non-routine analytical and non-routine interactive task inputs, which are associated with higher skill levels. Furthermore, it supports the hypothesis about the growing probability for women of being employed in an occupation that involve interpersonal skills (Cortes et al., 2018).

3.4 Changes in employment composition and wages

In Table 2 I present descriptive statistics on the evolution of the gender wage gap over time. Column (1) describes the changes for the period 1975 to 2010, while Columns (2) and (3) present two sub-periods: 1975 to 1992 and 1992 to 2010.

On average, between 1975 and 2010 there has been a reduction in the gender wage gap of 9.7 log points, which represents 21% of the gender wage gap in 1975. This convergence between male and female wages could be explained both, because of a decrease in the occupational segregation by gender, that is, women entering to better-paid occupations that were more male dominated in 1975, or because of a decrease in the gender gap in mean log wages in occupations in which women are represented. To understand the relevance of these two components I decompose the changes in the gender wage gap into a part explained by differential changes in employment across occupations for men ($\triangle s_{jt}^M$) and female workers ($\triangle s_{jt}^F$), taking the average wage between 1975 and 2010 in each occupation group j for male ($\bar{w}_{jt}^M$) and female workers ($\bar{w}_{jt}^F$): $\sum_j \left[\bar{w}_j^M \triangle s_{jt}^M - \bar{w}_j^F \triangle s_{jt}^F\right]$; and, on the other hand, a part explained by differential changes in male and female wages in each occupation, taking the average share between 1975 and 2010 of men and women in each occupation group: $\sum_j \left[\bar{s}_j^M \triangle w_{jt}^M - \bar{s}_j^F \triangle w_{jt}^F\right]$. The result of this decomposition is presented in Panel B.

I find that both channels contribute to explain the gender convergence across time. Around 70% of the reduction is explained by changes in average wages, while the other 30% is explained by changes in employment shares. The changes in average wages remain the most important component in both sub-periods, increasing its relevance to 86% in the period after 1992.

From another perspective, in Panel C, we can see that if only occupation shares had changed, keeping the wages of men and women in each occupation group at the values of 1975, the counterfactual gender wage gap in 2015 would be 41.5 (instead of 35.8), that is 16% higher. If instead only average wages in each occupation had changed, keeping the same distribution of men and women across occupations, the counterfactual gender wage gap in 2015 would be only 5% higher. That is, the most relevant change is when keeping average wages at the levels of 1975, since changes in wages rather than changes in average employment shares across occupation groups are the main cause of the the reduction in the gender gap during those 35 years.

With this background, in the next section I focus on discussing the results on the effect of technological change in wage changes by analyzing the estimated changes in the wage premiums and the role they play to explain the gender wage gap, given a certain composition of employment by gender in each year.

4 Results

4.1 Changes in wage premiums across occupations

I first estimate a regression where the changes in the occupation wage premiums ($\hat{\theta}_{jt}$) are identified through occupation-year fixed effects together for both genders. The changes in the wage premiums for each occupation reflect the evolution of wages for those who stay in the respective occupation. Figure 3 plots the estimated (mixed-by-gender) coefficients for the occupation-year dummies for a sample containing male and female workers in full-time jobs. Stars denote the level at which the estimated coefficients are significantly different from zero. As previously explained, these estimates should be interpreted as changes over time in the wage premium of each occupation relative to the base year and relative to the changes experienced by the base category (manual non-routine). In this case the base year is 1975.

The figure shows a relevant divergence in occupation premiums over time, and the patterns are consistent with the predictions of technological change hypothesis. Since the 1980s, the change in the occupation premiums for routine manual occupations evolve in a downward trend. For the cognitive occupations, the change in the wage premiums shows an upward trend, especially for those occupation groups that involve analytical and interactive non-routine tasks.

In Table 3 I analyze the contribution of the changes in the occupation premiums to explain the gender wage gap. The estimated changes in the wage premiums are reported for the last year with respect to the base year. Column (1) indicates the changes over time for the period 1975-2010, while Columns (2) and (3) decompose it for the two sub-periods. Panel A presents the results when using the mixed-by-gender estimates, weighted by the respective distribution shares of workers across occupations in the last year. We can observe that from 1975 to 2010 the average change in the occupation premiums increased for both, men and women and in each sub-period, with higher magnitudes for the more recent years. However, the increases for female workers are higher than those for male workers, leading to a negative change in the gender gap premia, that is, in favor of female workers. This implies that, absent the gender differences in the employment distribution across occupation groups the gender wage gap would have increased by about 11.6 log points, or 25% of the initial wage gap.

The fact that male workers are more affected by changes in the average wage premiums can be explained by an over-representation of men at manual-routine occupations, which exhibit decreasing wage premia over time for each cohort. That is, the gender differences in sorting across occupation acted as an equalizing force for the changes in the gender wage gaps.

4.2 Gender differences in occupation premiums

In this subsection I discuss the results when the wage premiums are allowed to differ by gender, that is, when estimating gender-specific changes in occupation premiums over time. These gender-specific wage premiums are estimated from equation (3), where an interaction between time-varying occupation fixed effects and a dummy for female workers captures the differential effect by gender. These estimates should be interpreted as changes over time in the wage premiums with respect to the variation in the manual non-routine for male workers.

Figure 4 plots the gender-specific changes over time in the occupational wage premiums for men and women separately. The results for men show an increasing trend in the occupation premiums from 1975 to 2010 for all cognitive occupations, with higher magnitudes for analytical non-routine occupations, and a significant decrease in the wage premium of manual routine occupations. These results are in line with those found by Cortes (2016) for male workers in the US. However, he considers manual and cognitive routine occupations together, and finds a decline in the wage premia for routine occupations overall. In this paper, I consider manual routine and cognitive routine occupations in two different task-based classification groups, and I show that the decreasing wage premia is mainly affecting manual routine occupations (an occupation group that represents around 52% of male employment), while the premium for cognitive routine occupations increases over this period, although less than for analytical and interactive non-routine occupations.

The changes in the occupation premiums estimated for female workers show a smaller dispersion between the different occupation groups than in the case of men, and they start to diverge later in time. For analytical non-routine occupations we can observe a significant upward trend since 1985. However, the trends for changes in the occupation premiums in interactive non-routine and cognitive routine occupations, in the case of women evolve similarly to the omitted category (manual non-routine for men), except for the last six years, when the premiums in analytical non-routine increase more rapidly. The large gender differences in the growing patterns of wage premiums for these two occupational groups is particularly relevant since these occupations represent more than half of the female labor force. This gender gap in the evolution of wage premiums might be partly explained by differences in composition within occupation groups (see Table C.2 in the Appendix). The most relevant occupations within Interactive non-routine groups for male workers are Entrepreneurs and managing directors (27.5%) and Foremen master mechanics (12.7%), which are supervisory positions, followed by Physicians (7.1%). For female workers, on the other hand the most representative occupations within the interactive group are Nursery teachers, child nurses (20.6%), Social workers, care workers (17.5%) and Home wardens, social work teachers (11%), which are feminized occupations related with care work. This aspect is further analyzed in Section 7.

Panel B in Table 3 shows the effect on the gender wage gap when using the gender-specific estimates for the occupation premiums. We can observe that when we allow occupation premiums to differ by gender, we obtain on average negative changes over time for females and positive for male workers (with respect to MNR for male workers). The gender gap in the change in occupation premiums rose by 18.2 log points between 1975 and 2010, with larger increases in the second part of the period. This result implies that absent the rise in occupation premiums inequality, the gender wage gap would have declined by 40% rather than the 21% we observe.

As previously discussed, the increase in the gender gap in occupation premiums could be explained by a widening on the differences in gender-specific occupation premiums (within occupation differences) or by the occupational segregation by gender (sorting across occupation groups), which implies that a different share of workers are affected by those relative changes across occupations. The results of the decomposition into these two possible channels is presented in the bottom part of Table 3. These results are based on the decomposition described by equation (5) which assigns women the occupation distribution of men, as this is a more relevant counterfactual. The findings for the alternative counterfactual (using female employment) are similar, and they are presented in Table A.1 in the Appendix.

We observe that the changes in the wage premiums that take place within occupation groups are the key drivers of the gender differences over time. That is, the fact that occupational wage premiums for men grow more rapidly than female premiums within certain occupation groups is the main explanatory factor of the average gender differences over time. The within occupation differences dominate as the main factor explaining the gender gap in occupation premiums for each period, and the effect is partially compensated by a negative contribution of the sorting across occupations channel. As previously stated, this could be explained by an over-representation of men at manual routine occupations (mainly industrial blue-collar occupations), which exhibit decreasing wage premiums over time for each cohort.

Both, the fact that the gender wage gap would decrease when using the occupation premiums that are common to both genders (instead of gender-specific premiums), and the dominating role of the within occupation channel in the decomposition of the gender differences, suggest that the sorting across occupations contributed to the narrowing of the gender gap, and point to a key role of growing within occupations wage premiums gender inequality, mainly in those occupations that involve cognitive non-routine tasks, to explain the persistence of a gender wage gap over time.

5 Cohort effects and changes over the work-life

In this section I explore for potential cohorts effects. The changes over time in the intensity of the introduction of technological progress might have affected differently male and female workers that entered the labor market at different periods. To investigate this, I consider four cohorts of male and female workers born between 1945 to 1965, and analyze their labor market trajectories during a period of twenty years.[20] In the next-subsection I discuss descriptive statistics on the evolution of the gender wage gaps across cohorts and over the work-life for each cohort of workers. Then I go on to present the results of estimating time-varying occupation premiums separately for each cohort and discuss its effects on the widening of the wage gaps over the work-life.

[20]For the cohort analysis the age intervals are restricted to 25-50 years old, so that the changes are calculated over the same number of years for each cohort. Cohort 1 includes individuals born in 1945-1950 and the changes consider period 1975-1995. Cohort 2 includes individuals born in 1951-1955 and the changes consider period 1980-2000. Cohort 3 includes individuals born in 1956-1960 and the changes consider period 1985-2005. Cohort 4 includes individuals born in 1961-1965 and the changes consider period 1990-2010.

5.1 Changes in the gender wage gaps across cohorts and over the work-life

Table 4 shows that the gender wage gap changed from being 45 log points on average for the cohort of workers born between 1945 and 1950 (column 1), to 36.5 log points (column 4) for the youngest cohort (born between 1961 and 1965). As it is possible to observe in Panel A, this reduction of gender inequality across cohorts is due to more equal wages between men and women at the starting point, when the individuals are between 25 and 30 years old rather than by a less steeper patter of wage gap growth over the work-life. In fact, the wage gaps greatly increase over the work-life, contrary to the small reduction over time previously discussed in the gender wage gap for a representative sample of workers. Figure 5 shows the evolution of the gender wage gap by age for each cohort. It seems that most of the decline in the gender wage gap through time is due to changes between cohorts. That is, younger cohorts show lower levels of the gender wage gap, specially before age 35. However, the shape of the work-life profiles, which indicates the changes in the gender wage gap within cohorts, shows similar patterns across cohorts, and a bit more steeper for the last cohort. That is, while the gap at entry has fallen, the evolution of the gap over the lifecycle increased.[21]

To analyze which part of these growing differences in wage trajectories of men and women over the work-life is explained by gender differential changes in employment across occupations and in the average wages by gender for each occupation, I perform the same decomposition as in Table 2, but this time comparing the year in which the individuals of each cohort are between 25 and 30 years old and the last, where they are between 45 and 50 years old. We can see that the difference in mean log wages between the last and first year is 16 log points for the first cohort and 25 log points for the youngest cohort. I find that these gender differences in the way in which wages grow through lifecycle are explained mostly by wage changes within occupation groups: between 73% and 75%. The changes in the share of men and women across occupation groups (considering wages at the work-life average for each gender and occupation), on the contrary, explain a smaller proportion (24% to 26%) of the increasing gender gap over the work-life.

This can be also observed in the analysis of the counterfactuals in Panel C. The increase in the gender wage gap over the work-life for each cohort is explained both by changes in average distribution across occupation groups and in average wages. Therefore, if we keep one of the components in its values of the base year, the gender wage gap would decrease, what explains the negative sign in the percentage difference for the counterfactual. If we consider the wages at the initial year and that only occupation groups composition changed over the work-life for each cohort, the counterfactual gender wage gap at 45-50 years old would be 28% lower for the

[21] Figure A.1 in the Appendix shows that the evolution of the wage trajectories over the work-life highly differs between occupations. The wage trajectories for cognitive occupations evolve in steeper patterns, specially in analytical and interactive non-routine occupations. While for both genders the increase in the wages is greater while younger, for women the growing tendency seems to stagnate after 32 years old. For men, on the contrary it stagnates for routine occupations but for cognitive non-routine occupations the wages continue increasing over the work-life. The slightly higher levels of wages for men at the beginning of the life-cycle, together with steeper wage trajectories, with men's wages growing faster than women's, specially in cognitive non-routine occupations is what explains a growing wage gap over the life-cycle. Therefore, changes in the average gender wage gap hide substantial heterogeneity in changes between cohorts as well as over the lifecycle.

oldest cohort and 47.6% lower for the youngest cohort. On the other hand, if only average wages changed, with the initial occupational composition of employment by gender, the counterfactual gender wage gap would be 12% lower for the oldest cohort and 23% lower for the youngest cohort.

5.2 Gender differences in changing occupation premiums over the work-life

I study first the changes in the wage premiums over the work-life when estimating coefficients that are common by gender. Figure 6 plots the changes over a period of 20 years in the occupational wage premium separately for each cohort.[22] The figures show similar trends, with decreasing occupation premiums for manual routine occupations and upward trends for cognitive occupations, especially so for analytical and interactive non-routine occupations. Comparing across cohorts we observe an increase in the dispersion of the changes in the occupation premiums, with steeper evolution trends in more recent cohorts compared to the oldest ones. These patterns suggest an increase in the wage premia inequality within cohorts. Those who most benefit from these changes are the workers that were born after 1955 and enter in the labor market around 1985 in analytical non-routine occupations, as they show the most steeper patterns in the wage premiums growth.

Table 5 Panel A presents the results for the gender wage gap when using these mixed-by-gender estimates, weighted by the respective distribution shares of workers across occupations in the last year for each cohort. We can observe that, during those 20 years the average change in the occupation premiums, increase for both men and women and for all the cohorts, with higher magnitudes for the more recent cohorts. However, the increases for female workers are higher than for male workers, leading to a negative change in the gender wage gap premia between -0.03 log points for the oldest cohort to -0.075 log points for the most recent cohort. Absent the gender differences in the employment distribution across occupation groups the gender wage gap would have increased by about 10% for the oldest cohort to 34% for the most recent one. That is, if women and men received the same occupation premiums, the gender wage gap would be closing over the work-life, and this difference in favor of female workers is accentuated for younger generations. The gender differences in sorting across occupation acted as an equalizing force for the changes in the gender wage gaps over the work-life.

To understand why this convergence did not occurred, I estimate gender-specific changes in the wage premiums across occupations for each cohort. These estimates are plotted in Figures 7 and 8. To help comparison between changes for male and female workers the estimates are plotted together for both genders, and separately for non-routine occupations (Figure 7) and for routine ones (Figure 8). For non-routine analytical and interactive occupations, I find that the wage premiums for men evolve in a much steeper pattern than those for females, and this difference is accentuated for the more recent cohorts. This means that the wage premium in cognitive non-routine occupations become more unequal across generations. In the case of routine occupations, we can observe a downward trend for manual routine occupations for both genders. However,

[22] The base years are 1975, 1980, 1985 and 1990 for each cohort respectively. In the base year the individuals of each cohort are aged between 25 and 30 years old.

the evolution of wage premiums for cognitive routine occupations shows a decreasing trend only for female workers, while it raises over time in the case of men.

In Table 5 the average gender gaps calculated with the gender-specific estimates rise from 9.5 log points for the oldest cohort to 17.7 log points for the most recent cohort. If we compare these estimates with the gender wage gap in the initial year, this result implies that the rise in occupation premiums inequality, leads to an increase of the gender wage gap over the work-life of 30% for the oldest cohort and 79% for the most recent.

As previously discussed, the increase in the gender gap in occupation premiums could be explained by a widening on the distribution of gender-specific occupation premiums (within occupation differences) or by the employment composition by gender across occupation groups (sorting across occupations). The results of the decomposition of the changes into these two possible channels for each cohort is presented in the bottom part of Table 5. We can see that the changes in the wage premiums that take place within occupation groups are the key drivers of the gender differences over time. That is, the fact that occupational wage premiums for men grow more rapidly than female premiums within the same occupations is the main explanatory factors of gender differences over time. The within occupation differences dominate the gender gap in occupation premiums for each cohort, and the effect is partially compensated by a negative contribution of the sorting across occupations channel.

6 Robustness analysis

This section presents a set of robustness checks on the empirical specification and analyses possible concerns that could affect the results.

6.1 Changing returns to education

The empirical strategy used so far assumes that the returns to ability do not vary with time. However, it is possible to extend it, to allow for changes over time in the return to some observable characteristics, which may affect ability. In particular, a possible concern would be that differences in the changes in the wage premium across occupation groups are driven by changes over time in the return to education. Previous work for West Germany find that the estimated returns to education for women greatly expanded in the period 1985-2002 (Ammermüller and Weber, 2005).

To address this issue, I follow Cortes (2016), assuming now that all individual skills (z_i) are fixed, but the return to education (a certain kind of observable skill) is allowed to vary over time: $\varphi_{jt}(z_i) = E_i\alpha_{jt} + \eta_i b_j$, where E_i captures the education level and η_i reflects all other individual abilities or skills. The return to these other skills is still assumed to be time-invariant, while the return to education varies over time but, for simplicity, it is assumed to be the same for all occupations ($\alpha_{jt} = \alpha_j + \alpha_t$). In this case I perform two separate regressions for men and women.[23] The regression to be estimated for the potential wage of individual i of gender g at

[23] Due to the number of variables introduced in the regression it is not possible to have all the interactions

year t is:

$$w_{it}^g = \sum_j D_{ijt}\theta_{jt}^g + E_i\alpha_t^g + \sum_j D_{ijt}v_{ij}^g + X_{it}\delta^g + u_{it} \tag{7}$$

where $g = \{M, F\}$ and $v_{ij}^g = E_i\alpha_j^g + \eta_i b_j^g$. The occupation spell estimated through this regression will now contain the return to education in the base year and the return to unobserved ability, but not the changes in the returns to education over time. I consider the three educational levels described in Section 3: i) lower secondary, intermediate secondary or upper secondary school leaving certificate, ii) at least a vocational qualification, and iii) a degree from a university of applied sciences or a university degree.

The estimation results show that there has been a growing pattern in the returns to having a university degree. For men it took place since 1985 (Figure A.2), while for women this has been so especially since the 90s (Figure A.3). Figures A.4 and A.5 show the estimated occupation effects using the baseline specification (*panel a*) and the new specification with occupation effects which are ridden off of the time-varying returns to education (*panel b*). While the most outstanding features of both figures remain unchanged, it is possible to observe that the new estimated wage premiums for analytical and interactive non-routine occupations in *panel b* are lower than those of *panel a*. This means that a portion of the growing patterns in the wage premiums in these occupations are explained by raising returns to education. This is quite expected as these are occupations that demand high educated profiles and the skill premiums grew during this period. The most notable change is that, after accounting for the effect of changing returns to education the wage premium in the interactive non-routine occupations is lower than that of cognitive routine. Changing returns to education do not seem to play a role for wage premiums in routine manual occupations, as it remains relative unchanged, with a clear downward trend for men, and similar to that of the manual non routine group (the omitted category) for women. Analyzed in the framework of the polarization patterns, changing returns to education seem relevant to explain the higher wage premiums in the upper part of the occupation distributions (analytical and interactive non-routine) but not the decreasing wage premiums in the middle part (routine occupations).

6.2 Occupation-specific tenure profiles

Another concern with the occupation premiums previously estimated might be the existence of occupation-specific human capital in the different occupation categories, that led to heterogeneous returns to occupational tenure (see Kambourov and Manovskii (2009) and Gathmann and Schönberg (2010)). If for example the tenure profile is more steeper in non-routine analytical and interactive occupations than in routine ones, this could affect my finding that the occupation wage premiums in manual routine occupations is falling over time.

To estimate wage premiums controlling for the effect of occupation-specific tenure profiles, I

with gender for all occupations and years. In this case the change in the wage premiums for men and women are gender-specific but no longer comparable given that the base is different.

follow the approach of Cortes (2016), introducing a return to individual's occupational tenure. More specifically, I estimate the following regression:

$$w_{it} = \sum_j D_{ijt}\theta_{jt} + \sum_j D_{ijt}\beta_{jt}\, fem_i + \sum_j D_{ijt}\gamma_{ij} + \sum_j D_{ijt}F_j(Ten_{ijt}) + X_{it}\delta + u_{it} \tag{8}$$

where Ten_{ijt} is individual's i's tenure in occupation j at time t and $F_j(Ten_{ijt})$ is a non-linear function that captures the occupation-specific returns to tenure. In this setting an individual's occupational choice will depend, as before, on his/her skill level z_i and the occupation wage premium θ_{jt}, but also on his/her tenure in the current occupation Ten_{ijt}. Sorting across occupations is still determined by skill cutoffs, but now these skill cutoffs would differ for individuals with different levels of occupational tenure.

I use a quadratic function of occupational tenure, interacted with occupation dummies to allow for different returns to tenure across the broad occupation groups. To define the occupational tenure I consider the days that the individual has been employed in a certain occupation group. I use as reference a variable that indicates the time that the individual has been in a certain firm, and I start occupational tenure from zero each time the individual switches his/her occupation group. Time out of the labor force and time in unemployment, as well time in apprenticeship training, is not counted. Also, if an employee returns to his occupation, after being out of the labor force or in another occupation I start counting occupational tenure from zero.

The gender-specific wage premiums estimated from equation (8) are presented in Figure A.6 in the Appendix. We can observe that they are very similar to those previously presented in Figure 4, with growing wage premiums over time in analytical and interactive non-routine occupations and decreasing premiums over time in manual routine occupations. The only noticeable difference is that the trend in wage premiums for manual non-routine occupations is less steeper than in Figure 4. That is, the general patterns in the changes in wage premiums across occupations remain robust to controlling for heterogeneity in tenure profiles.

6.3 Changes in selection into employment and attrition

A possible concern is that the patterns we observe are being driven by changing characteristics of working women (and men) over time (Mulligan and Rubinstein, 2008). That is, that the changes in the average occupation premiums reflect the influence of how the "quality" of workers within these occupations change. Indeed, if we analyze the composition of the workforce, it is possible to observe that it becomes more educated (Table A.3), and this is true both for male and female workers. Therefore it is possible that the estimated changes in the occupation premiums in Figure 4 are affected by changes in the composition of the workforce over time (more positive selection into employment). However, the growing patterns in the occupation premiums in analytical and interactive non-routine occupations, as well as the decreasing premium that I find for manual routine occupations also hold when estimating changes in the occupation premiums separately for each cohort, and we can argue that they are less likely to reflect changes in the skill levels of the workforce, given that after 25 years old most of them already completed their education cycle. That is, selection issues become less relevant when comparing changes in occupation premiums for men and women within cohorts.

However, since I am observing changes in a period of twenty years for each cohort, a potential concern is that nonrandom attrition in terms of which individuals are still observed at longer time horizons may be biasing the estimations for the changes in wage premiums. This would be the case for example if many women leave the labor market after 30 or 35 years old due to maternity, and the characteristics of these women are different than those who remain employed within each occupation group. To address this concern I run the regressions for the estimation of the time-varying occupation premiums imposing the condition that the individuals are observed employed in at least 16 years (however, the patterns do not change if I impose the same sample of workers in all years - 20 years instead of 16). The resulting estimates are presented in Figures A.7 and A.8 for non-routine cognitive and routine occupation groups respectively, and are very similar to the estimated changes in the occupation premiums previously discussed in Figures 7 and 8. The only noticeable change is that for the most recent cohort the estimates for wage premium in analytical non-routine occupations for female workers are a bit higher when restricting the sample controlling for attrition, and also for cognitive routine the wage premiums show a less decreasing trend. We can therefore conclude that the results are not driven by differential attrition within occupation groups.

7 Motherhood and gender segregation within occupation groups

The evidence collected so far suggests that gender differences in the increasing dispersion of occupation premiums play an important role in explaining the persistence of the gender wage gap over time. This section investigates possible explanations behind the gender differences in wage premiums growth over time.

7.1 Motherhood

A recent literature has provided evidence for high-income countries that motherhood drives sizable and persistent gaps in earnings, employment rates and hours of work (Adda et al., 2017; Angelov et al., 2016; Kleven et al., 2019). In this subsection I investigate in which way childbirth is related with the flatter gender-specific wage premiums trajectories for women, by estimating changes in the occupation premium over the work-life for mothers and non-mothers.

Although child birth dates are not directly recorded in my data, there is information on the reason of cancellation/notification of an employment (or unemployment) spell, which allows me to identify child-birth related work interruptions and estimate the date of birth.[24] I focus on mothers working full-time before and after child-birth. This might generate a selection bias since

[24] More details can be found in Müller and Strauch (2017). A caveat of this approach is that child-births that take place before women enter the labor market cannot be identified, and that successive births are often hard to detect. However, this last problem is not a relevant issue here, as I focus on the first child, that is, when the woman becomes a mother.

many women return to the labor market with reduced working hours. However, since there is no information on the hours of work, keeping only those working full-time is necessary to make wage premiums comparable among them.

To estimate wage premiums for mothers and non-mothers I follow the same approach that I used for estimating gender-specific occupation premiums, but this time keeping only the sample of female workers and introducing an interaction of occupation-year fixed effects with a dummy that takes value one for mothers. Therefore, the omitted category is the routine non-manual group for non-mothers, and the estimated coefficients should be interpreted as changes over time in the occupational wage premium relative to the base year and relative to the changes experienced by the manual non-routine for non-mothers.

In my sample of women, I identify 162,829 individuals who are mothers and 283,783 non-mothers. The average age at birth is 29.7 years old. Figures 9 and 10 plot the estimated parameters for the changes in the wage premiums over the work-life for each cohort of mothers and non-mothers employed in non-routine and in routine occupations respectively. The results show a clearly differentiated pattern in the wage premiums trajectories of mothers and non-mothers. While the wage premiums in analytical and interactive non-routine occupations for non-mothers tends to increase over the working life, for mothers they show a decreasing trajectory during the first ten years. There are also relevant differences between cohorts. For younger cohorts the wage premia trajectory of mothers shows a inverted-U shape, with increasing returns in the last years. Also, wage premiums differences across occupation groups tend to increase for younger cohorts. In Figure 9 it is possible to observe that for non-mothers the premium in analytical non-routine occupations increases more rapidly than that of interactive occupations, and this is more pronounced the younger the cohort is. Similarly, in Figure 10 the wage premiums in cognitive routine occupations for non-mothers show an increasing trajectory across working life which becomes more steeper for younger cohorts. However, this is not verified for mothers, as the wage premiums in cognitive routine occupations in the case of mothers is the one which verifies the largest decreases.

Table 6 presents the changes in the occupation premiums over the work-life averaged using the specific occupation distribution of employment across occupation groups of mothers and non-mothers in the last year. Mothers obtain on average negative changes in their occupation premiums over the work-life, while non-mothers experience positive changes. After twenty years the difference in the wage premiums for non-mothers and mothers is between 20 and 27 log points.[25]

The bottom part of the table shows the decomposition of the difference in occupation premiums into sorting across occupations and within occupation differences. The changes in occupation premiums that take place within occupation groups, that is, the fact that the occupation premiums for non-mothers grow more rapidly than those of mothers within the same occupation groups, is the main driver of the differences between mother and non-mothers over time. The sorting across occupations channel has a positive effect for all cohorts except for the oldest

[25] If we include also those with part time jobs the difference in the wage premiums for non-mothers and mothers is between 27 and 36 log points.

one, that is, non-mothers are more represented in occupations with increasing wage premiums, which contributes to wide the wage premium gap between mothers and non-mothers over time, although it explains less than 3% of the increase.

An interesting question is how the dynamics for women with no children evolve relative to those of men.[26] To analyze this I run two different regressions with the group of men and mothers and non-mothers separately. Figure A.9 and A.10 in the Appendix plot the changes in the wage premiums of men and women with no child (left panel) and of men and mothers (right panel) pooling all cohorts together. In Figure A.9 we can observe that the wage premiums in analytical and interactive non routine occupations present a growing pattern for both men and childless women, however, the levels are lower for childless women compared to men. On the contrary, for mothers the wage premium shows a decreasing trajectory in analytical and interactive occupations compared to men. Regarding the changes in the wage premiums in routine occupations, Figure A.10 shows that for cognitive routine occupations presents a growing trajectory for both men and childless women, with higher levels of growth for men. On the other hand, for manual routine occupations it presents a decreasing trajectory, with larger declines in the wage premiums for men than for childless women. However, this is not the case for mothers, where the wage premiums in both cognitive and manual routine occupations shows decreasing patterns with larger declines for mothers than for men. These results suggest that child-related explanations might be one of the factors behind the lower wage premiums growth over the work-life that female workers experience in comparison with male workers.

7.2 Gender segregation within non-routine cognitive occupations

From the estimation of gender-specific occupation premiums, I find that gender differences within occupational groups are the main explanation for the slowdown of gender wage convergence. In particular, wage premiums for analytical and interactive non-routine occupations increase much more rapidly for men than for women. In order to understand the sources of these gender differences, in this section I investigate more deeply the wage trajectories of men and women, by looking at the composition of the detailed occupations in which men and women are employed within those two occupation groups.

Predominantly feminized occupations pay less than those with a lower share of women, even after adjusting for education and skills (Levanon et al., 2009). The literature has provided two different hypothesis to explain it. One explanation is that the pay offered in an occupation affects the proportion of female workers that it employs. A second view is the "devaluation hypothesis", which states that women's entry into a male-dominated occupation diminish the value of this occupation and, consequently, its relative wage (England et al., 2000). Therefore, the proportion of female workers in an occupation affects its wages. The devaluation hypothesis is expected to affect mostly some high-paid male dominated occupations, which absorbed more women in recent decades. This might be the case in certain interactive and analytic non routine occupations, that increase the percentage of women in large numbers.

[26] It is not possible to identify fathers with these data.

I focus on male and female workers employed in analytical and interactive non-routine occupations, and consider the five most relevant disagregated occupations at the 3 digit level of the German Classification of Occupations 1988 (KldB 1988) in each of these two occupational groups (Table 7). For females the most relevant occupations in the analytical group are Technical draughtsperson (representing 15.9% of the employed in that group), Data processing specialists (15.7%), Other technicians (9.5%), Economic and Social scientists (5.9%) and Chemical laboratory assistants (5.9%). In total these five single occupations represent 53% of the women employed in this group. However, all these are occupations in which the proportion of male workers is higher than that of females. In all these occupations the gender wage gap is higher than 30 log points, except for Chemical laboratory assistants with an average gender wage gap of 25 log points.

Within the interactive non-routine group the most relevant occupations for female workers are Nursery teachers, child nurses (17.6%), Social workers, care workers (13.7%) and Home wardens, social work teachers (9.4%). These are feminized occupations: female workers represent 96%, 77% and 65% of the employed in these three occupations respectively. They also have relatively lower gender wage gaps (the difference in log wages is between 0.14 and 0.21 log points) relative to the other occupations within the interactive group in which men are more represented. Other occupations that are relevant for women within the interactive group are Entrepreneurs, managing directors (7.9%) and Physicians (5.2%). On average over the period, in these two occupations women represent 16.9% and 37.4% of the workers respectively, but are occupations that have experienced a considerable growth in the number of women employed. Also, they show large gender wage gaps, specially for Entrepreneurs, managing directors, that has a gender wage gap of 0.62 log points. For male workers the most relevant occupations within the interactive non-routine group are Entrepreneurs, managing directors (29.5%), Foremen master mechanics (13.7%), which are supervisory positions, followed by Physicians (7.4%).

This evidence suggest that part of the gender differences in the wage trajectories within analytical and interactive non-routine groups are due to the gender segregation within these groups. In the interactive non-routine group, women are more represented in feminized occupations, where the wage trajectories are flatter than those in which men are mostly employed. On the other hand, in those occupations in which women are newcomers, they face lower wage increases than male workers, which expands the gender gap over time.

8 Conclusions

Driven by technological advances, labor markets in high-income countries have witnessed relevant structural changes in the relocation of employment from manual routine cognitive work to non-routine cognitive tasks. This paper uses administrative panel data for West Germany to investigate the effect of technological change on the dynamic of gender wage differentials. I find that gender segregation across occupations has mostly benefited female workers, contributing to narrow the gender gap. This is mostly explained by the fact that men are more exposed to automation of work than women, given their over representation in manual routine occupations (mainly industrial blue-collar jobs), but also because women increased their employment in cognitive non-routine occupations. However, by investigating male and female wage trajectories across occupation groups, I find that the wage gains for male workers within cognitive occupations grew more rapidly than those of females, with the effect being strongest for most recent cohorts. Then, my findings suggest that although women have been less exposed to the automation of work and increased their employment in non-routine high-pay occupations, they still face certain constrains that did not allow them to benefit from the increased overall wage returns in the upper part of the skill distribution.

More research is needed to understand the underlying reasons behind the larger growth of male premiums relative to those of females within analytical and interactive non-routine occupations. Nevertheless, I provide some suggestive evidence that two factors might play a role. By looking at the more detailed occupational composition I present evidence showing that one of the reasons behind the differences in wage gains is that women moved disproportionately to interactive occupations and within this group, to occupations that were already highly feminized, such as nursery teachers and social workers, which experienced lower wage growth over time compared to occupations where male workers relocated and which are likely to be characterized by greater complementarities between labor and technology. Also, I provide some evidence showing that women show flatter wage trajectories than men due to having children, which would be a factor explaining the difference in wage gains over time in cognitive non-routine occupations. This is an aspect which is not related to a different effect of or exposure to technological change, but to other restrictions that women might face in the labor market, such as the fact that they are still the main carer for children and the way in which certain occupations are structured and remunerated (Goldin, 2014), among other institutional factors.

Within the questions left unanswered by this analysis, future research will investigate further the individual-level adjustments of male and female workers in the labor market and the differential changes in wages that men and women experience both as stayers in a certain occupation group and when switching out from routine occupations, and how wage gains vary depending on the direction of the switch. Another channel to explore is the role of firms in explaining different wage gains for men and women within certain occupations. A recent literature has shown the relevance of firm-level wage differentials in explaining gender gaps. In particular, for Germany, Bruns (2019) estimates that gender differences in firm premiums explain around 15% of the gender wage gap between the 1990s and 2000s. The different sorting of men and women across firms within the broad occupation groups could lead to different impact of technological change for example, because women tend to be disproportionately employed in small firms where the introduction of technology is less likely. However, this might be also explained by the fact that men and women sort into different detailed occupations within the broad occupation groups, as previously shown, and this gender segregation might be also taking place within firms. Further research in those lines would help us to better understand the mechanisms through which technological change can differently affect male and female workers. Additionally, it would provide evidence that can help in the design of policies to remove obstacles for women's performance in the labor market and to reduce gender wage inequalities.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

Ammermüller, A. and A. M. Weber (2005). Educational attainment and returns to education in Germany. ZEW-Centre for European Economic Research Discussion Paper 05-017.

Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics 34*(3), 545–579.

Antonczyk, D., B. Fitzenberger, and U. Leuschner (2009). Can a task-based approach explain the recent changes in the german wage structure? *Jahrbücher für Nationalökonomie und Statistik 229*(2-3), 214–238.

Antoni, M., A. Ganzer, P. vom Berge, et al. (2016). Sample of integrated labour market biographies (siab) 1975-2014. Technical report, Research Data Centre (FDZ). IAB. FDZ-Datenreport 4.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in us wage inequality: Revising the revisionists. *The Review of Economics and Statistics 90*(2), 300–323.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Azmat, G. and B. Petrongolo (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics 30*, 32–40.

Bachmann, R., M. Cim, and C. Green (2019). Long-run patterns of labour market polarization: Evidence from German micro data. *British Journal of Industrial Relations 57*(2), 350–376.

Black, S. E. and A. Spitz-Oener (2010). Explaining women's success: technological change and the skill content of women's work. *The Review of Economics and Statistics 92*(1), 187–194.

Blau, F. D. and L. M. Kahn (2017). The gender-wage gap: Extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources 8*(4), 436–455.

Borghans, L., B. Ter Weel, and B. A. Weinberg (2014). People skills and the labor-market outcomes of underrepresented groups. *ILR Review 67*(2), 287–334.

Bruns, B. (2019). Changes in workplace heterogeneity and how they widen the gender wage gap. *American Economic Journal: Applied Economics 11*(2), 74–113.

Büttner, T. and S. Rässler (2008). Multiple imputation of right-censored wages in the german iab employment sample considering heteroscedasticity. Technical report, Research Data Centre (FDZ). IAB- Discussion Paper 44/2008.

Card, D., A. R. Cardoso, P. Kline, et al. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics 131*(2), 633–686.

Cortes, G. M. (2016). Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics 34*(1), 63–105.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2018). The "end of men" and rise of women in the high-skilled labor market. National Bureau of Economic Research.

Cortes, G. M., A. Oliveira, and A. Salomons (2020). Do technological advances reduce the gender wage gap? *Oxford Review of Economic Policy 36*(4), 903–924.

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

England, P., J. Hermsen, and D. Cotter (2000, 05). The devaluation of women's work: A comment on tam. *American Journal of Sociology 105*, 1741–1751.

Fitzenberger, B., A. Osikominu, and R. Völter (2005). Imputation rules to improve the education variable in the iab employment subsample. Technical report, Research Data Centre (FDZ). IAB- Discussion Paper 05/2005.

Gartner, H. and S. Rässler (2005). Analyzing the changing gender wage gap based on multiply imputed right censored wages. Technical report, Research Data Centre (FDZ). IAB- Discussion Paper 05/2005.

Gathmann, C. and U. Schönberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics 28*(1), 1–49.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of labor economics 23*(4), 681–724.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-Economic Panel Study (SOEP). *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics 239*(2), 345–360.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review 104*(4), 1091–1119.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of american college women: The reversal of the college gender gap. *The Journal of Economic Perspectives 20*(4), 133–133.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in Europe. *American Economic Review 99*(2), 58–63.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review 104*(8), 2509–26.

Kambourov, G. and I. Manovskii (2009). Occupational specificity of human capital. *International Economic Review 50*(1), 63–115.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Levanon, A., P. England, and P. Allison (2009). Occupational feminization and pay: Assessing causal dynamics using 1950-2000 us census data. *Social Forces 88*(2), 865–891.

Müller, D. and K. Strauch (2017). Identifying mothers in administrative data. Technical report, Research Data Centre (FDZ). IAB- Discussion Paper 13/2017.

Mulligan, C. B. and Y. Rubinstein (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics 123*(3), 1061–1110.

Ngai, L. R. and B. Petrongolo (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics 9*(4), 1–44.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review 14*(3), 693–709.

Olivetti, C. and B. Petrongolo (2014). Gender gaps across countries and skills: Demand, supply and the industry structure. *Review of Economic Dynamics 17*(4), 842–859.

Olivetti, C. and B. Petrongolo (2017). The economic consequences of family policies: Lessons from a century of legislation in high-income countries. *The Journal of Economic Perspectives 31*(1), 205–230.

Petrongolo, B. and M. Ronchi (2020). Gender gaps and the structure of local labor markets. *Labour Economics 64*, 101–819.

Rohrbach-Schmidt, D. and M. Tiemann (2013). Changes in workplace tasks in germany. evaluating skill and task measures. *Journal for Labour Market Research 46*(3), 215–237.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics 24*(2), 235–270.

Tables

Table 1: Summary statistics

	All sample				Full-time			
	Men		Women		Men		Women	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
A. Worker characteristics								
Age	39.36	8.59	39.26	8.80	39.42	8.58	38.26	8.97
Part-time	0.02	0.15	0.33	0.47	0.00	0.00	0.00	0.00
Experience	11.63	7.94	11.91	8.30	11.64	7.92	11.10	7.94
German nationality	0.89	0.31	0.92	0.26	0.89	0.31	0.91	0.28
Real daily wage	107.9	137.6	62.1	77.0	109.4	138.7	72.4	88.5
Low Skill	0.10	0.30	0.15	0.36	0.10	0.30	0.15	0.36
Medium Skill	0.76	0.43	0.75	0.44	0.76	0.43	0.74	0.44
High Skill	0.15	0.35	0.11	0.31	0.14	0.35	0.11	0.31
B. Occupation classification								
Analytical NR	0.14	0.34	0.04	0.20	0.14	0.34	0.05	0.21
Interactive NR	0.08	0.27	0.12	0.32	0.07	0.26	0.11	0.32
Cognitive R	0.13	0.33	0.39	0.49	0.13	0.33	0.41	0.49
Manual R	0.52	0.50	0.14	0.35	0.53	0.50	0.17	0.37
Manual NR	0.14	0.34	0.31	0.46	0.13	0.34	0.26	0.44
N of individuals	565,230		446,612		565,230		446,611	
Observations	7,102,028		4,761,753		6,924,032		3,199,340	

Notes: Data corresponds to SIAB, period 1975 - 2010. The sample is composed of male and female workers, aged between 25 and 55 years old in West Germany, excluding apprentices and marginal part-time employees. Real daily wages are expressed in euros of 2010.

Table 2: Changes in the gender wage gap

	1975-2010 (1)	1975-1992 (2)	1992-2010 (3)
Av. gender gap (mean log wages)	0.399	0.432	0.370
A.Wage gap changes			
Gender gap in the base year	0.455	0.455	0.398
Change across the period	-0.097	-0.057	-0.040
Gender gap in the final year	0.358	0.398	0.358
B. Decomposition of the change			
Changes in occupation distribution	-0.029	-0.019	-0.005
% of the gap	*30.05*	*33.63*	*13.52*
Changes in average wages	-0.068	-0.038	-0.035
% of the gap	*69.95*	*66.37*	*86.48*
C. Counterfactuals: gap in final year			
If only occupation shares had changed	0.415	0.432	0.389
% difference	*15.99*	*8.60*	*8.84*
If only average wages had changed	0.376	0.413	0.360
% difference	*5.13*	*3.90*	*0.64*

Notes: Data corresponds to SIAB, period 1975 - 2010. Only workers in full-time jobs are included. The gender gaps are expressed in mean log daily real wages.

Table 3: Decomposition of the changes in the gender wage premiums gap

	1975-2010 (1)	1975-1992 (2)	1992-2010 (3)
Gender gap in the base year	0.455	0.455	0.398
Change in gender gap in mean log wages	-0.097	-0.057	-0.040
A. Mixed by gender occupation wage premiums			
$E[\theta_{jt}^{ALL}\|male]$	0.175	0.046	0.129
$E[\theta_{jt}^{ALL}\|female]$	0.290	0.081	0.210
$E[\theta_{jt}^{ALL}\|male] - E[\theta_{jt}^{ALL}\|female]$	-0.116	-0.035	-0.080
% of the initial wage gap	*-25.39*	*-7.71*	*-20.22*
B. Gender-specific occupation wage premiums			
$E[\theta_{jt}^{M}\|male]$	0.161	0.041	0.119
$E[\theta_{jt}^{F}\|female]$	-0.022	-0.001	-0.020
$E[\theta_{jt}^{M}\|male] - E[\theta_{jt}^{F}\|female]$	0.182	0.043	0.140
% of the initial wage gap	*40.06*	*9.37*	*35.10*
C. Decomposition			
Sorting across occupations			
$E[\theta_{jt}^{F}\|male] - E[\theta_{jt}^{F}\|female]$	-0.021	-0.003	-0.018
% of the gender difference in occup premiums	*-11.47*	*-5.94*	*-13.16*
Within occupation differences			
$E[\theta_{jt}^{M} - \theta_{jt}^{F}\|male]$	0.203	0.045	0.158
% of the gender difference in occup premiums	*111.47*	*105.94*	*113.16*
Observations in last year	278,142	310,952	278,142

Notes: Data corresponds to SIAB, period 1975 - 2010. Only workers in full-time jobs are included. $E[\theta_{jt}^{ALL}|male]$ is the mean value of the mixed by gender occupation premia across male workers (considering male's distribution of occupations in the last year), while $E[\theta_{jt}^{ALL}|female]$ takes the mean across female workers. $E[\theta_{jt}^{M}|male]$ and $E[\theta_{jt}^{F}|female]$ are the mean values of male and female specific occupation premia across male and female workers respectively.

Table 4: Changes in the gender wage gap over the work-life

	Cohort 1 (1)	Cohort 2 (2)	Cohort 3 (3)	Cohort 4 (4)
Av. gender gap (mean log wages)	0.451	0.405	0.380	0.365
A.Wage gap changes				
Gender gap in the base year	0.319	0.273	0.228	0.223
Change across the period	0.164	0.203	0.226	0.248
Gender gap in the final year	0.483	0.476	0.454	0.471
B. Decomposition of the change				
Changes in occupation distribution	0.042	0.050	0.059	0.066
% of the gap	*25.67*	*24.68*	*26.11*	*26.49*
Changes in average wages	0.122	0.153	0.167	0.182
% of the gap	*74.33*	*75.32*	*73.89*	*73.51*
C. Counterfactuals: gap in final year				
If only occupation shares had changed	0.347	0.296	0.254	0.246
% difference	*-28.24*	*-37.71*	*-44.04*	*-47.62*
If only average wages had changed	0.427	0.399	0.362	0.363
% difference	*-11.68*	*-16.06*	*-20.25*	*-22.84*

Notes: Data corresponds to SIAB, period 1975 - 2010. Only workers in full-time jobs are included. The gender gaps are expressed in mean log daily real wages. Cohort 1 includes individuals born in 1945-1950 and the changes consider period 1975-1995, while they are between 25-30 to 45-50 years old. Cohort 2 includes individuals born in 1951-1955 and the changes consider period 1980-2000. Cohort 3 includes individuals born in 1956-1960 and the changes consider period 1985-2005. Cohort 4 includes individuals born in 1961-1965 and the changes consider period 1990-2010.

Table 5: Decomposition of the changes in the wage premiums gap over the work-life

	Cohort 1 (1)	Cohort 2 (2)	Cohort 3 (3)	Cohort 4 (4)
Gender gap in the base year	0.319	0.273	0.228	0.223
Change in gender gap in mean log wages	0.164	0.203	0.226	0.248
A. Mixed by gender occupation wage premiums				
$E[\theta_{jt}^{ALL}\|male]$	0.068	0.091	0.099	0.118
$E[\theta_{jt}^{ALL}\|female]$	0.101	0.137	0.145	0.193
$E[\theta_{jt}^{ALL}\|male] - E[\theta_{jt}^{ALL}\|female]$	-0.033	-0.047	-0.046	-0.075
% of the initial wage gap	*-10.34*	*-17.06*	*-20.31*	*-33.78*
B. Gender-specific occupation wage premiums				
$E[\theta_{jt}^{M}\|male]$	0.068	0.099	0.100	0.119
$E[\theta_{jt}^{F}\|female]$	-0.028	-0.047	-0.068	-0.058
$E[\theta_{jt}^{M}\|male] - E[\theta_{jt}^{F}\|female]$	0.095	0.147	0.168	0.177
% of the initial wage gap	*29.96*	*53.80*	*73.51*	*79.38*
C. Decomposition				
Sorting across occupations				
$E[\theta_{jt}^{F}\|male] - E[\theta_{jt}^{F}\|female]$	-0.010	-0.005	-0.008	-0.026
% of the gender difference in occup premiums	*-10.20*	*-3.25*	*-4.80*	*-14.67*
Within occupation differences				
$E[\theta_{jt}^{M} - \theta_{jt}^{F}\|male]$	0.105	0.152	0.176	0.203
% of the gender difference in occup premiums	*110.20*	*103.25*	*104.80*	*114.66*
Observations in last year	44,757	42,634	46,648	53,884

Notes: Data corresponds to SIAB, period 1975 - 2010. Only workers in full-time jobs are included. Cohort 1 includes individuals born in 1945-1950 and are analyzed during period 1975-1995. Cohort 2 includes individuals born in 1951-1955 and are analyzed in 1980-2000. Cohort 3 includes individuals born in 1956-1960 and are analyzed in 1985-2005. Cohort 4 includes individuals born in 1961-1965 and are analyzed in 1990-2010. Individuals in each cohort are followed while they are between 25-30 to 45-50 years old. $E[\theta_{jt}^{ALL}|male]$ is the mean value of the mixed by gender occupation premia across male workers (considering male's distribution of occupations in the last year), while $E[\theta_{jt}^{ALL}|female]$ takes the mean across female workers. $E[\theta_{jt}^{M}|male]$ and $E[\theta_{jt}^{F}|female]$ are the mean values of male and female specific occupation premia across male and female workers respectively.

Table 6: Decomposition of the changes in the gender wage premiums gap between mothers and non-mothers

	Cohort 1 (1)	Cohort 2 (2)	Cohort 3 (3)	Cohort 4 (4)
Gender gap in the base year	-0.068	0.002	0.038	0.017
Change in gender gap in mean log wages	0.152	0.134	0.168	0.285
Mother/non-mother wage premiums				
$E[\theta_{jt}^{non} \mid non-mother]$	0.049	0.054	0.077	0.123
$E[\theta_{jt}^{mom} \mid mother]$	-0.146	-0.107	-0.114	-0.141
$E[\theta_{jt}^{non} \mid non-mother] - E[\theta_{jt}^{mom} \mid mother]$	0.196	0.162	0.191	0.265
Decomposition				
Sorting across occupations				
$E[\theta_{jt}^{mom} \mid non-mother] - E[\theta_{jt}^{mom} \mid mother]$	-0.001	0.002	0.005	0.005
% of the difference in wage premium	*-0.73*	*0.99*	*2.80*	*1.93*
Within occupation differences				
$E[\theta_{jt}^{non} - \theta_{jt}^{mom} \mid non-mother]$	0.197	0.160	0.186	0.259
% of the difference in wage premium	*100.73*	*99.01*	*97.20*	*98.07*
Observations in last year	14,425	14,159	14,753	16,559

Notes: Data corresponds to SIAB, period 1975 - 2010. Only women in full-time jobs are included. Cohort 1 includes individuals born in 1945-1950 and are analyzed during period 1975-1995. Cohort 2 includes individuals born in 1951-1955 and are analyzed in 1980-2000. Cohort 3 includes individuals born in 1956-1960 and are analyzed in 1985-2005. Cohort 4 includes individuals born in 1961-1965 and are analyzed in 1990-2010. Individuals in each cohort are followed while they are between 25-30 to 45-50 years old.

Table 7: Most relevant occupations within analytical and interactive groups

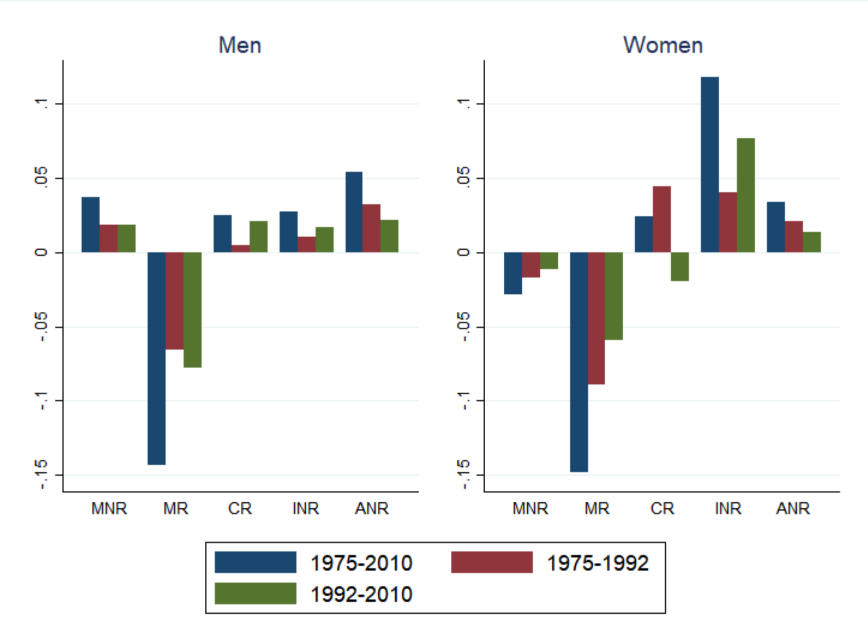
A. Analytical non-routine	N	% in group	% females	Var 75-2010	Mean wage	Wage at age 55	Gender wage gap
Women							
Technical draughtsperson	23,646	15.94	45.25	14.13	4.23	4.31	0.32
Data processing specialists	23,302	15.71	17.90	3.27	4.72	4.68	0.30
Other technicians	14,076	9.49	11.59	3.72	4.38	4.40	0.41
Economic and social scientists	8,727	5.88	32.25	25.73	4.74	5.19	0.38
Chemical laboratory assistants	8,679	5.85	38.77	20.96	4.45	4.59	0.25
Dental technicians	7,814	5.27	38.66	17.64	4.06	3.93	0.38
Men							
Other technicians	137,037	14.46	11.59	3.72	4.79	4.86	0.41
Data processing specialists	134,584	14.20	17.90	3.27	5.02	5.27	0.30
Electrical engineers	79,557	8.40	3.88	3.98	5.28	5.54	0.24
Electrical engineering technicians	75,236	7.94	3.90	2.79	4.79	5.01	0.31
Mechanical, motor engineers	74,597	7.87	2.98	5.05	5.27	5.49	0.38

B Interactive non-routine	N	% in group	% females	Var 75-2010	Mean wage	Wage at age 55	Gender wage gap
Women							
Nursery teachers, child nurses	62,803	17.60	96.32	-4.54	4.28	4.44	0.14
Social workers, care workers	48,900	13.70	77.02	5.61	4.23	4.28	0.21
Home wardens, social work teachers	33,488	9.38	65.55	1.90	4.38	4.53	0.17
Entrepreneurs, managing directors	28,053	7.86	16.94	5.85	4.53	4.57	0.62
Masseurs, physiotherapists	21,239	5.95	66.83	14.33	4.04	4.18	0.21
Physicians	18,435	5.17	37.42	22.29	5.10	5.49	0.29
Men							
Entrepreneurs, managing directors	152,297	29.46	16.94	5.85	5.16	5.27	0.62
Foremen, master mechanicals	70,638	13.67	2.20	1.89	4.82	4.89	0.44
Physicians	38,371	7.42	37.42	22.29	5.39	5.81	0.29
Managements consultants, organisers	26,806	5.19	23.38	22.66	5.25	5.38	0.31
Home wardens, social work teachers	24,062	4.65	65.55	1.90	4.54	4.71	0.17

Notes: Data corresponds to SIAB, period 1975 - 2010. Only workers in full-time jobs are included.

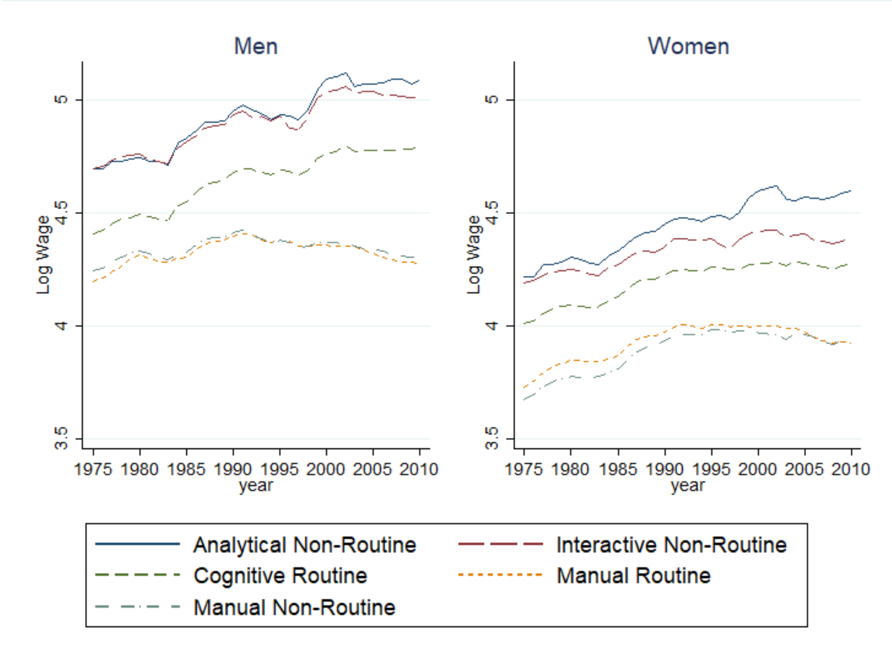
Figures

Figure 1: Changes in employment shares by occupation groups



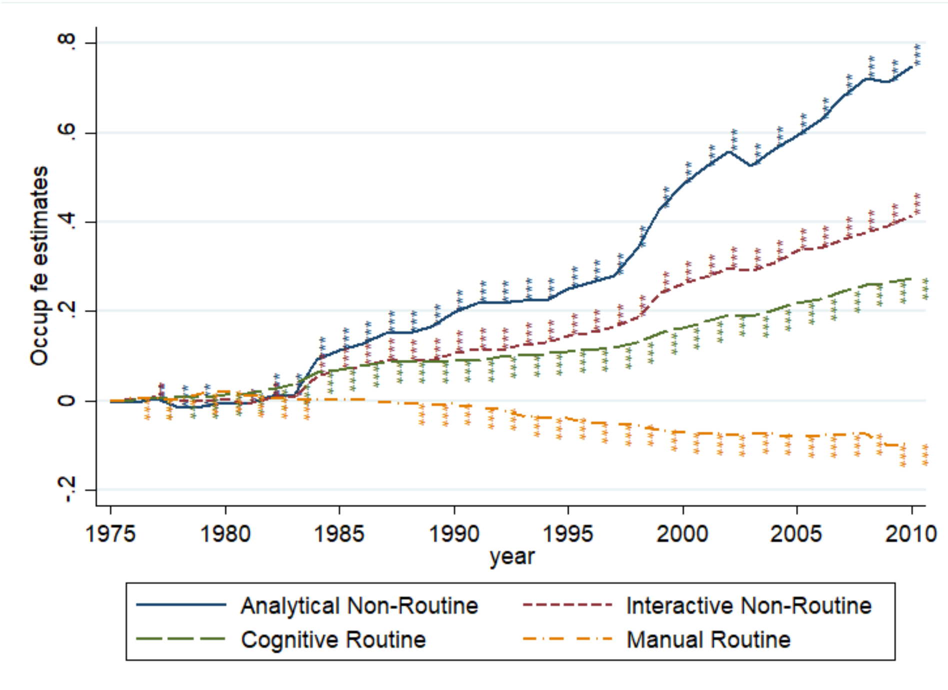
Notes: Task groups are ranked by their average wage. The five occupational groups considered are Manual non-routine (MNR), Manual routine (MR), Cognitive routine (CR), Interactive non-routine (INR), and Analytical non-routine (ANR). See Section 3 for details on the construction of this classification of occupational groups.

Figure 2: Evolution of log real daily wage for broad occupation groups (1975-2010)



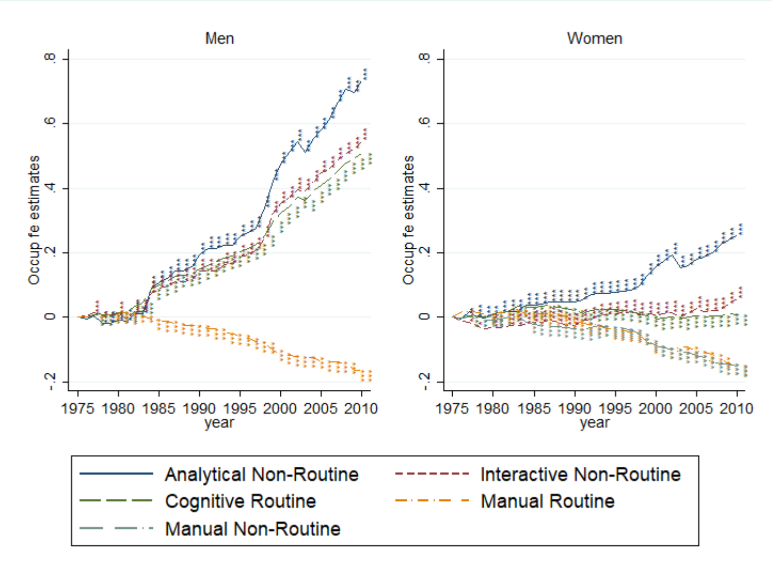
Notes: Only observations corresponding to full-time workers are included.

Figure 3: Change in occupational premiums. Mixed-by-gender fixed effects



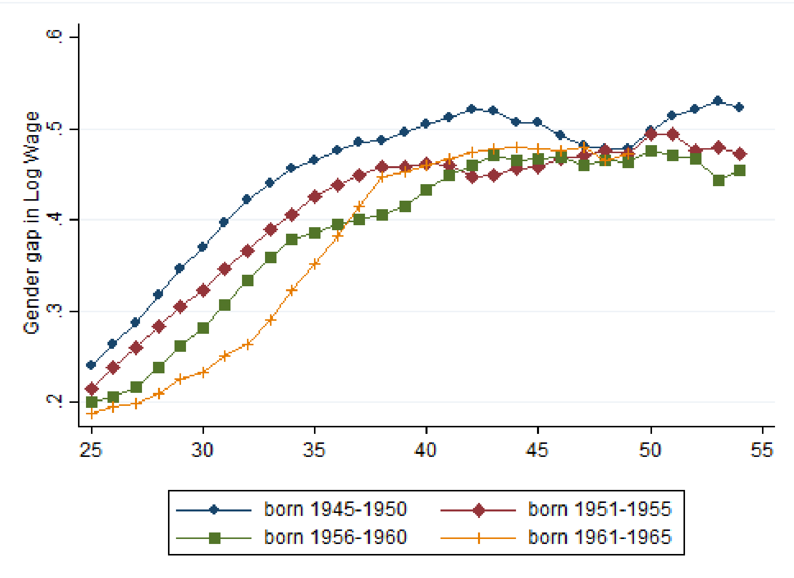
Notes: the figure plots the estimated coefficients on occupation-year dummies for a sample containing both male and female workers in full-time jobs. Stars denote the level at which the estimated coefficients are significantly different from zero (*** p<0.01, ** p<0.05, * p<0.1)

Figure 4: Change in gender-specific occupational premiums



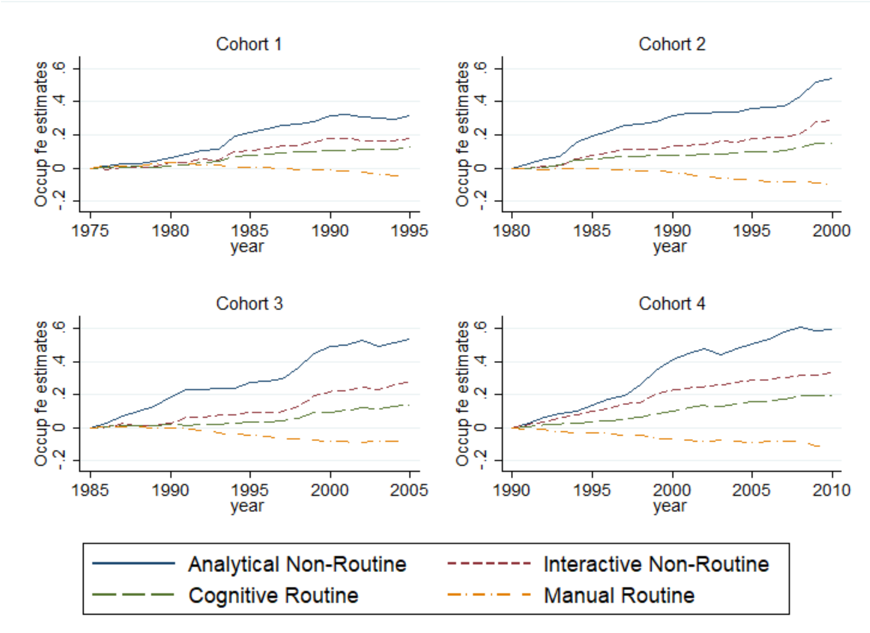
Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers in full-time jobs. Stars denote the level at which the estimated coefficients are significantly different from zero (*** p<0.01, ** p<0.05, * p<0.1). They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

Figure 5: Gender gap in mean of Log daily real wage by age for each cohort



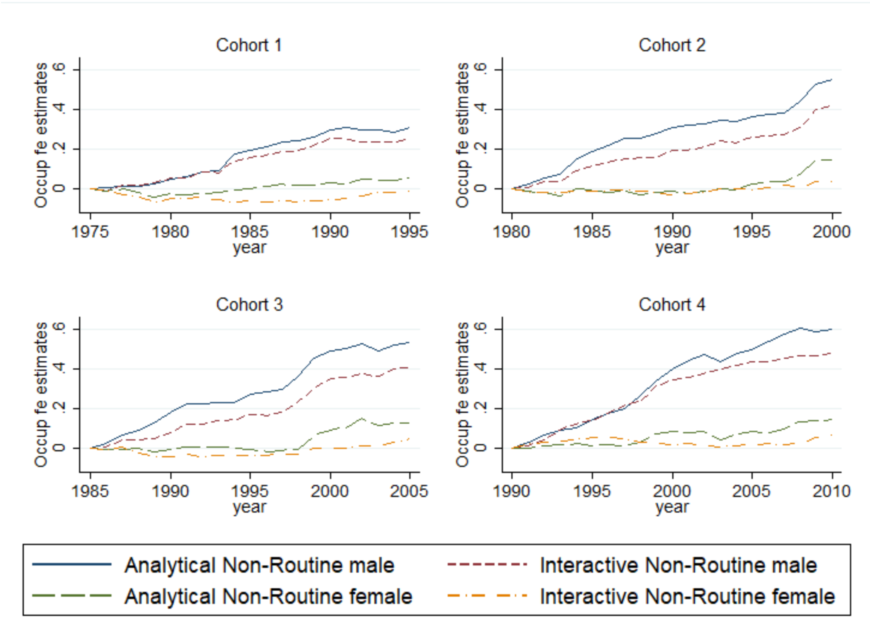
Notes: The graphs plot means of the gender gap of log real daily wages for each cohort by age. Only observations corresponding to full-time workers are included.

Figure 6: Change in occupational premiums by cohorts. Mixed-by-gender fixed effects



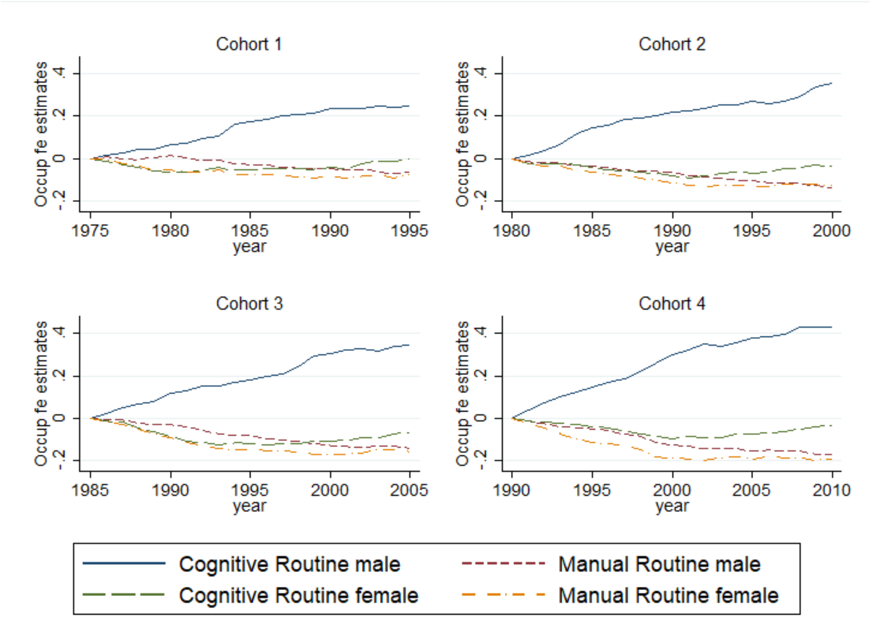
Notes: the figure plots the estimated coefficients on occupation-year dummies separately for each cohort for a sample containing both male and female workers in full-time jobs. Stars were omitted to help visualization.

Figure 7: Change in gender-specific occupational premiums. Non-routine cognitive occupations



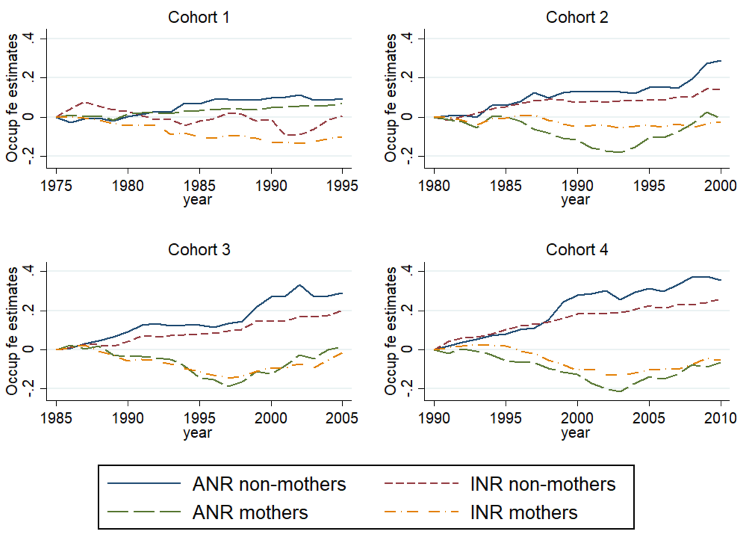
Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers in full-time jobs in analytical and interactive non-routine occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men. Stars were omitted to help visualization.

Figure 8: Change in gender-specific occupational premiums. Routine occupations



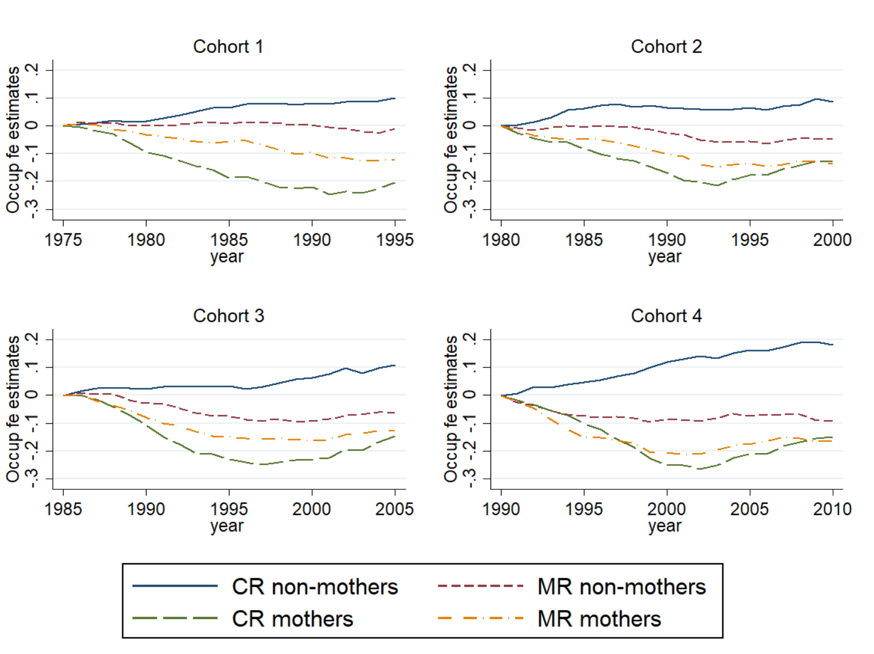
Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers in full-time jobs in cognitive routine and manual routine occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men. Stars were omitted to help visualization.

Figure 9: Change in occupational premia for mothers and non-mothers by cohort. Non-routine cognitive occupations



Notes: The two occupation groups included in this figure are Analytical Non-Routine (ANR) and Interactive Non-Routine (INR). The figure plots the estimated coefficients on occupation-year dummies for mothers and non-mothers in full-time jobs. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for non-mothers.

Figure 10: Change in occupational premia for mothers and non-mothers by cohort. Routine occupations



Notes: The two occupation groups included in this figure are Cognitive Routine (CR) and Manual Routine (MR). The figure plots the estimated coefficients on occupation-year dummies for mothers and non-mothers in full-time jobs. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for non-mothers.

APPENDIX

A Additional results

A.1 Tables

Table A.1: Alternative decomposition of the changes in the gender gap of occupation premiums

	1975-2010 (1)	1975-1992 (2)	1992-2010 (3)
Gender gap in the base year	0.455	0.455	0.398
Change in gender gap in mean log wages	-0.097	-0.057	-0.040
Gender-specific occupation wage premiums			
$E[\theta^M_{jt}\|male] - E[\theta^F_{jt}\|female]$	0.182	0.043	0.140
% of the initial wage gap	*40.06*	*9.37*	*35.10*
Decomposition			
Sorting across occupations			
$E[\theta^M_{jt}\|male] - E[\theta^M_{jt}\|female]$	-0.174	-0.049	-0.124
% of the gender difference in occup premiums	*-95.21*	*-115.31*	*-89.07*
Within occupation differences			
$E[\theta^M_{jt} - \theta^F_{jt}\|female]$	0.356	0.092	0.264
% of the gender difference in occup premiums	*195.21*	*215.31*	*189.07*
Observations in last year	278,142	310,952	278,142

Notes: Only workers in full-time jobs are included. $E[\theta^M_{jt}|males]$ and $E[\theta^F_{jt}|females]$ are the mean values of male and female occupation premiums across male and female workers respectively.

Table A.2: Alternative decomposition of the changes in the wage premiums gap over the work-life

	Cohort 1 (1)	Cohort 2 (2)	Cohort 3 (3)	Cohort 4 (4)
Gender gap in the base year	0.319	0.273	0.228	0.223
Change in gender gap in mean log wages	0.164	0.203	0.226	0.248
Gender-specific occupation wage premiums				
$E[\theta_{jt}^M \vert male] - E[\theta_{jt}^F \vert female]$	0.095	0.147	0.168	0.177
% of the initial wage gap	*29.96*	*53.80*	*73.51*	*79.38*
Decomposition				
Sorting across occupations				
$E[\theta_{jt}^M \vert male] - E[\theta_{jt}^M \vert female]$	-0.058	-0.093	-0.103	-0.13544
% of the gender difference in occup premiums	*-60.39*	*-63.20*	*-61.60*	*-76.67*
Within occupation differences				
$E[\theta_{jt}^M - \theta_{jt}^F \vert female]$	0.153	0.240	0.271	0.312092
% of the gender difference in occup premiums	*160.39*	*163.20*	*161.60*	*176.66*
Observations in last year				

Notes: Only workers in full-time jobs are included. Cohort 1 includes individuals born in 1945-1950 and are analyzed during period 1975-1995. Cohort 2 includes individuals born in 1951-1955 and are analyzed in 1980-2000. Cohort 3 includes individuals born in 1956-1960 and are analyzed in 1985-2005. Cohort 4 includes individuals born in 1961-1965 and are analyzed in 1990-2010. Individuals in each cohort are followed while they are between 25-30 to 45-50 years old. $E[\theta_{jt}^M|males]$ and $E[\theta_{jt}^F|females]$ are the mean values of male and female occupation premiums across male and female workers respectively.

Table A.3: Educational levels composition by cohorts

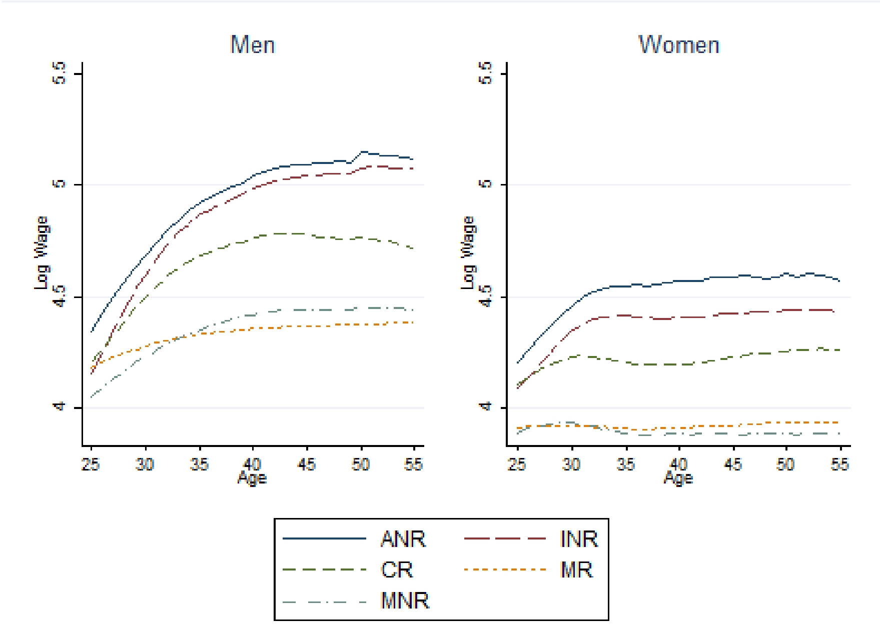
Men				
	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Low	9.36	6.48	6.37	6.58
Medium	77.70	77.87	77.12	75.92
High	12.95	15.65	16.51	17.51
N observ	903,370	879,544	944,286	920,722

Women				
	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Low	17.96	12.04	8.83	7.60
Medium	76.18	78.22	78.25	77.99
High	5.86	9.74	12.93	14.42
N observ	405,524	431,544	451,317	436,471

Notes: Only workers in full-time jobs are included. The low skill level comprise individuals with a lower secondary, intermediate secondary or upper secondary school leaving certificate but no vocational qualifications. The medium skill level include those with a lower secondary, intermediate secondary or upper secondary school leaving certificate and a vocational qualification. The high skilled include all employees who have a degree from a university of applied sciences (*Fachhochschule*), technical college degree or a university degree.

A.2 Figures

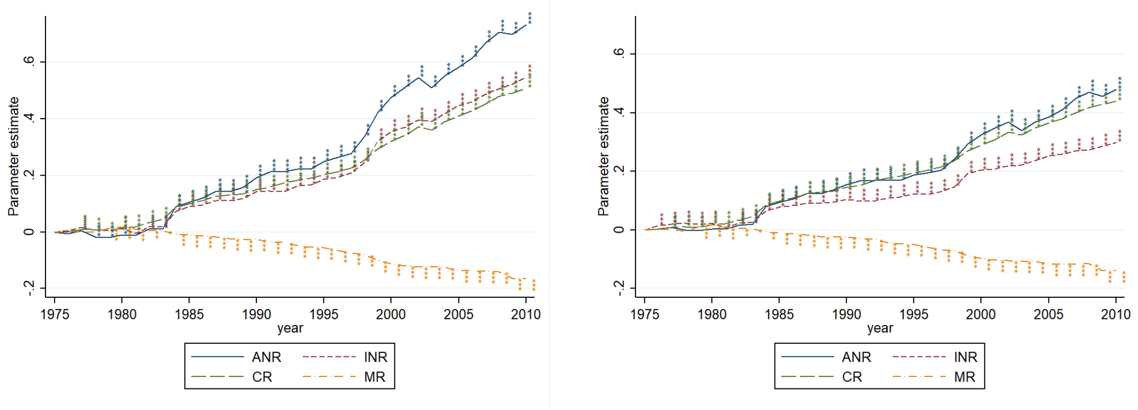
Figure A.1: Mean of Log daily real wage by age for each broad occupation group



Notes: The graphs plot means of the logarithms of real daily wages in the five occupational groups by age. Only observations corresponding to full-time workers are included.

Figure A.2: Change in occupational premiums controlling for time-varying returns to education. Male workers

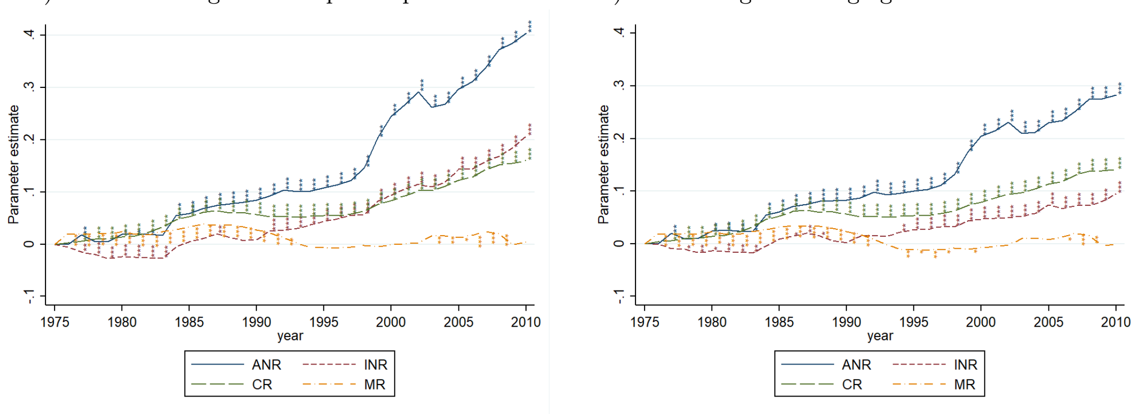
a) Baseline changes in occupation premiums b) Controlling for changing returns to education



Notes: the figure plots the estimated coefficients in occupation-year dummies for men and women with no children (left) and for men and mothers (right) in full-time jobs in analytical and interactive occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

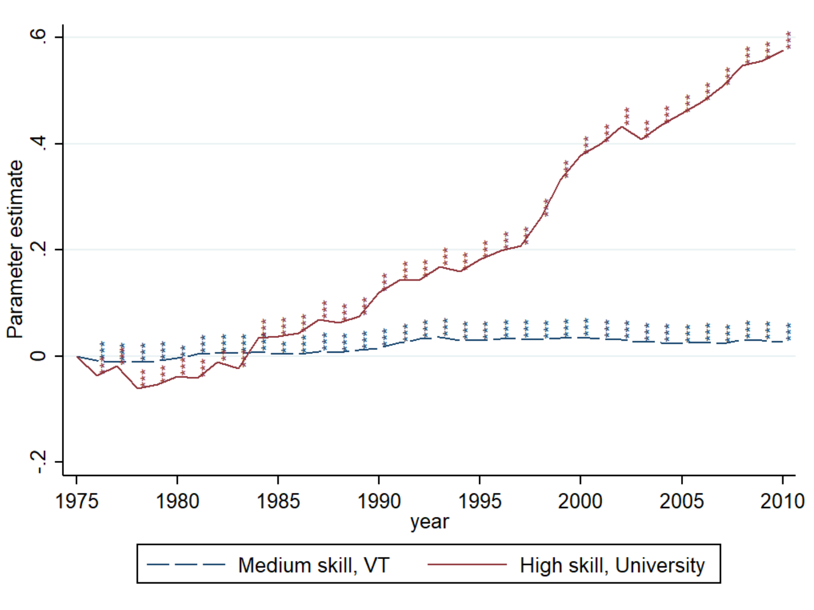
Figure A.3: Change in occupational premiums controlling for time-varying returns to education. Female workers

a) Baseline changes in occupation premiums b) Controlling for changing returns to education



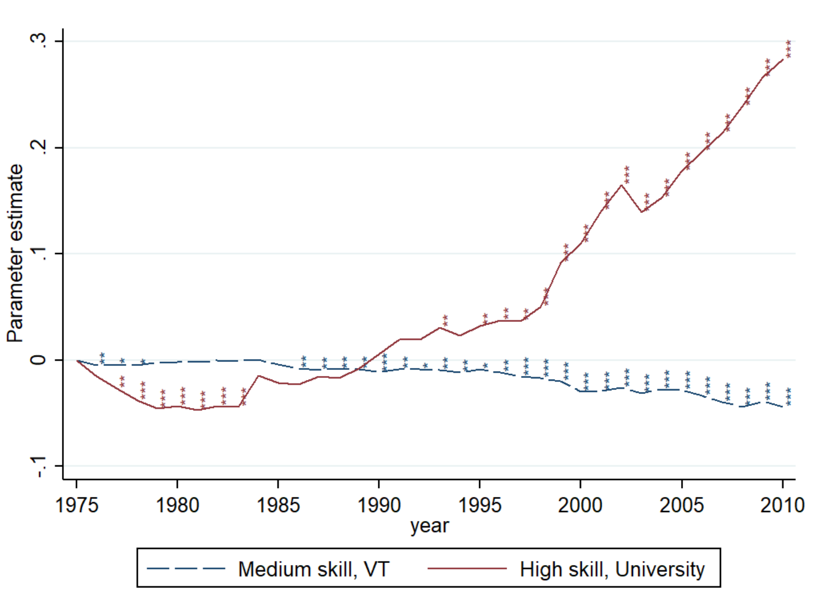
Notes: the figure plots the estimated coefficients in occupation-year dummies for men and women with no children (left) and for men and mothers (right) in full-time jobs in analytical and interactive occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

Figure A.4: Changes in the returns to education over time. Male workers



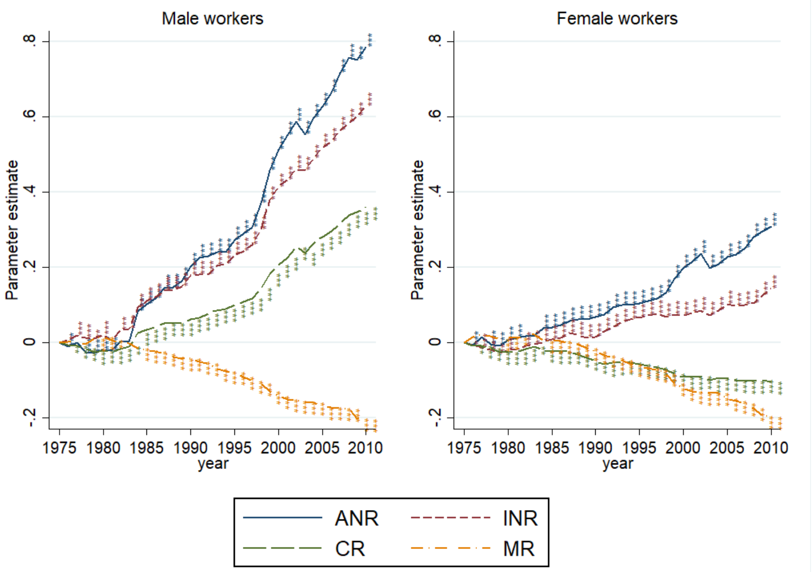
Notes: the figure plots the estimated coefficients in education-year dummies for men in full-time jobs.

Figure A.5: Changes in the returns to education over time. Female workers



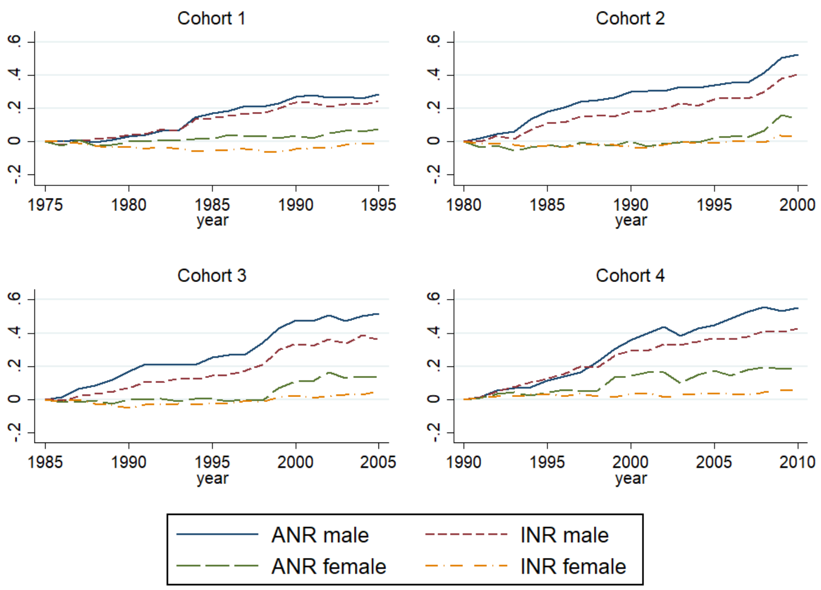
Notes: the figure plots the estimated coefficients in education-year dummies for women in full-time jobs.

Figure A.6: Change in occupational premiums controlling for heterogeneous occupation tenure profiles



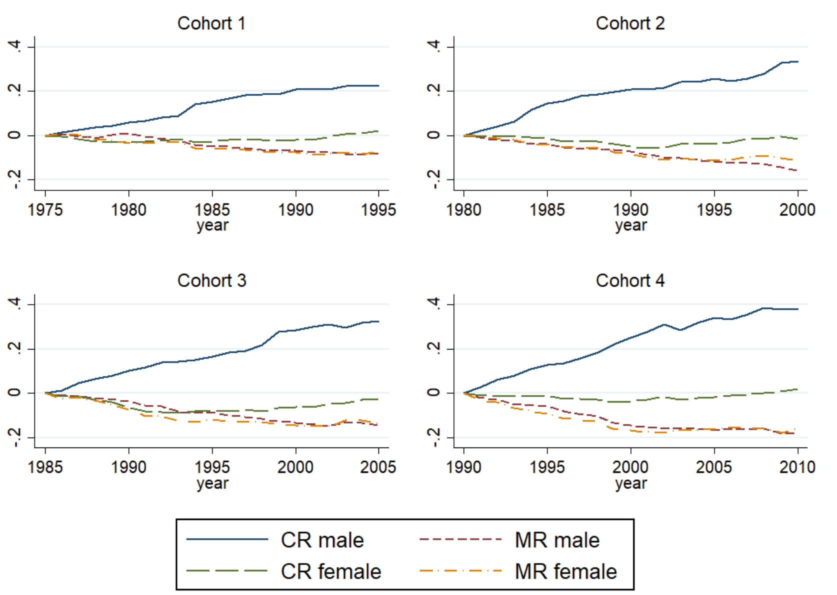
Notes: the figure plots the estimated coefficients in occupation-year dummies for men and women with no children (left) and for men and mothers (right) in full-time jobs in analytical and interactive occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

Figure A.7: Change in gender-specific occupational premiums controlling for attrition. Non-routine cognitive occupations



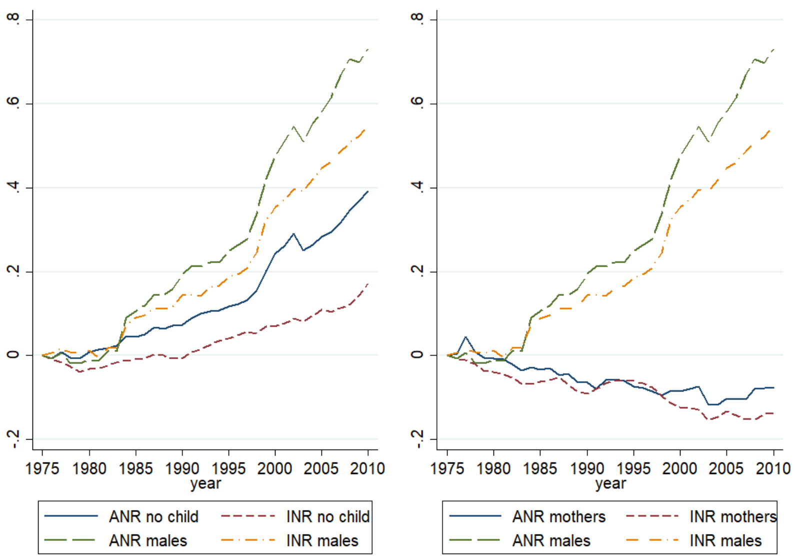
Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers in full-time jobs. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men. Stars were omitted for better visualization.

Figure A.8: Change in gender-specific occupational premiums controlling for attrition. Routine occupations



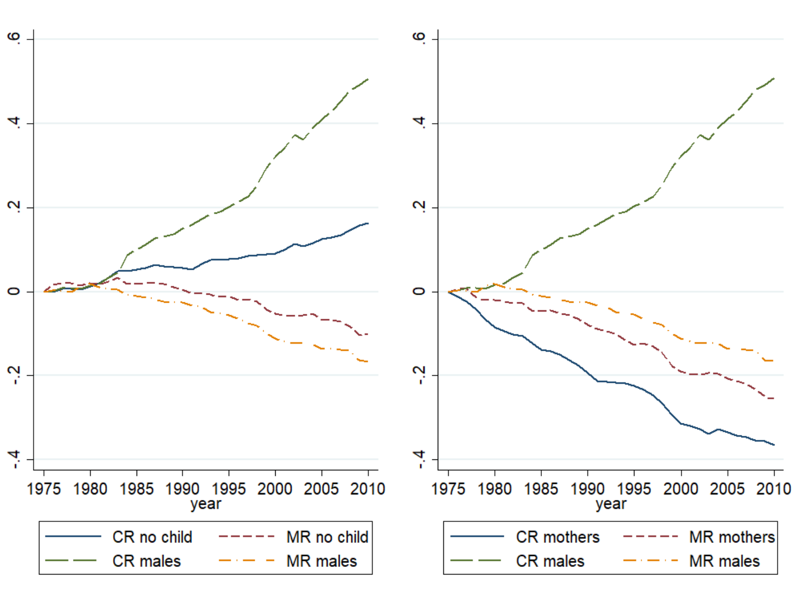
Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers in full-time jobs in cognitive routine and manual routine occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men. Stars were omitted for better visualization.

Figure A.9: Change in occupational premiums mothers and non-mothers compared to men. Cognitive non-routine occupations



Notes: the figure plots the estimated coefficients in occupation-year dummies for men and women with no children (left) and for men and mothers (right) in full-time jobs in analytical and interactive occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

Figure A.10: Change in occupational premiums mothers and non-mothers compared to men. Routine occupations



Notes: the figure plots the estimated coefficients in occupation-year dummies for men and women with no children (left) and for men and mothers (right) in full-time jobs in cognitive routine and manual routine occupations. They measure the change relative to 1975 in the occupation premia with respect to the analogous change of manual non-routine for men.

B Extended sample including part-time workers

The empirical strategy used in the paper restricts the analysis to workers employed in full-time jobs in order to make wages comparable across workers. This is because the SIAB dataset has one drawback: it does not contain information on hours worked, and therefore does not allow to construct an hourly wage measure. Excluding part-time workers might be a concern given the fact that part-time work is quite frequent among female workers in Germany.

To address this issue, I conduct a robustness check using data on hours worked from an alternative dataset: the German Socio-Economic Panel Study (SOEP-Core). The SOEP study is a wide-ranging annual representative longitudinal survey of private households, conducted by the German Institute for Economic Research, DIW Berlin. The data provide information on all household members, reaching around 15,000 households and about 30,000 persons, living in the Eastern and Western German States from 1984 to 2018 (see Goebel et al. (2019) for further information). These data contain information on actual work hours per week, which allows me to estimate average number of hours worked by part-time and full-time workers.

In order to ensure consistency with my estimation sample in the SIAB, I restrict the SOEP data to workers in West Germany, aged 25 to 55 years old and excluding apprentices and marginal part-time workers, but including regular part-time workers. For this sample I estimate the mean hours worked separately for men and women in part-time and full-time jobs for each year from 1984 to 2010.[27] In this sample in the SOEP data the part-time workers represent 17.7% of the workers. For male workers 2.5% of the observations correspond to part-time while this percentage rises to 40.8% for female workers. The average hours worked for full-time workers are 44.5 hours for male workers and 41.4 for female workers. For part-time workers the average hours are 26.4 for males and 22.7 for female workers.[28]

I impute the average hours worked by gender and year to part-time and full-time workers in the SIAB data, and construct an approximately hourly wage, dividing the administrative information on individual's daily wages by the average hours worked (divided by seven to convert from hours per week to daily hours worked). Then I use this measure of hourly wages to estimate wage premiums across occupation groups following the empirical strategy described in Section 2, keeping now a sample that contains both, full-time and part-time workers, and where the dependent variable in the regression is hourly wage instead of daily wage.

In my sample of the SIAB data the part-time workers represent 16% of the observations. Among male workers 3% of the observations are part-time, while among female workers 34.4% are part-time. Focusing on the broad occupation groups, the highest percentage of part-time for men takes place in interactive non-routine occupations, where it represents 8.3% of the employed in this group. For women the highest percentage of part-time is observed in manual non-routine occupations (44.8%), followed by interactive non-routine (36.48).

Table B.1 presents the means of logarithms of hourly wages calculated in the SOEP data by dividing the monthly labor income by the individual's declared hours of work (multiplied by 4) and the means of logarithms of hourly wages in SIAB data obtained by dividing the administrative information of individual's daily wages by the average hours of work by individuals of same sex and type of job (part-time/full-time) in each year. Except for female workers in part-time jobs, the means obtained in SOEP data are slightly higher than those approximated for the administrative data.

Then I estimate the changes in the occupation premiums using hourly wages and including both part-time and full-time workers. For comparison, I present also the estimated changes in wage premiums using daily wages and only full-time workers considering only the period 1984-2010, that is when I have information on hours worked, so that the reference year for the estimated changes is the same.

I first estimate the regression where the changes in the occupation wage premiums are identified through occupation-year fixed effects together for both genders. Figure B.1 plots the estimated (mixed-by-gender) coefficients for the occupation-year dummies. On the left hand side the figure shows the estimated occupation premiums using daily wages for a sample containing male and female workers in full-time jobs, while in the right hand side, the sample includes also part-time workers, and the occupation premiums were estimated using hourly wages. As previ-

[27] Corresponding weighting factors are used to calculate the means.

[28] Given the reduced number of observations for male workers in part-time jobs, it is not possible to take means of hours worked including other characteristics such as education, region or age groups. If I do so, the number of observations in each year for each cell used to compute the means would be below the one requested by IAB for an external aggregate database.

ously explained, these estimates should be interpreted as changes over time in the wage premium of each occupation relative to the base year and relative to the changes experienced by the base category (manual non-routine). In this case the base year is 1984. As it is possible to observe, the patterns in both figures are very similar. Consistent with technological change hypothesis, they show a downward trend in occupation premiums for routine manual occupations, while for the cognitive occupations the change in the wage premiums shows an upward trend, especially for those occupation groups that involve analytical and interactive non-routine tasks. The only noticeable difference is that the trends including part-time workers show a slightly less steep pattern of growth, particularly for interactive non-routine and cognitive-routine occupations, which are the occupations where most of part-time workers are employed.

Figures B.2 shows the results of estimating gender-specific changes in occupation premiums over time for the sample of full-time workers only (Panel A) and when including also part-time and using hourly wages (Panel B). Comparing the figure where only full-time workers are included with the one that includes also part-time, for male workers there is almost no change. For female workers, on the other hand, including part-time workers shows a slightly higher upward trend compared with the ones including only full-time workers, notably in the changes in occupation premiums for the manual non-routine and interactive non-routine occupations. For interactive non-routine occupations, when including part-time workers the trends of change in occupation premiums are more close to those of analytical, and for some particular years, even larger. In spite of these differences, we can state that the main features of the changes in gender-specific wage premiums across occupations remain unchanged for both, male and female workers.

In Table B.2 I analyze the contribution of the changes in the occupation premiums to explain the gender wage gap. The estimated changes in the wage premiums are reported for the last year with respect to the base year. Column (1) indicates the estimated changes over time for the period 1984-2010 using daily wages for the sample of full-time workers (as in the main analysis in the paper), Column (2) presents the estimated changes for full-time workers but using the approximate measure of hourly wages, while Column (3) includes both full-time and part-time workers. Panel A presents the results when using the mixed-by-gender estimates, weighted by the respective distribution shares of workers across occupations in the last year. We can observe that from 1984 to 2010 the average change in the occupation premiums increased for both, men and women. However, the increases for female workers are higher than for male workers, leading to a negative change in the gender gap premia, that is, in favor of female workers. This implies that, absent the gender differences in the employment distribution across occupation groups the gender wage gap would have increased by about 9.8 log points (10 log points when using hourly wages), or 22% of the initial wage gap for full-time workers and 9 log points or 24% of the initial wage gap for the sample including also part-time workers.

Panel B shows the effect on the gender wage gap when using the gender-specific estimates for the occupation premiums. For the sample of full-time workers, we can observe that when we allow occupation premiums to differ by gender, we obtain on average negative changes over time for females and positive for male workers (with respect to MNR for male workers). For the sample including part-time workers the changes over time in wage premiums are positive for both male and female workers, but higher for the males. The gender gap in the change in occupation premiums between 1984 and 2010 rose by 15 log points for full-time workers (14 log points when using hourly wages), and by 10.4 log points for the sample including both part-time and full-time workers. As before, the results when including part-time workers show that gender differences in sorting across occupations acted as an equalizing force for the changes in the gender wage gaps. This would be explained by an over-representation of male workers in manual-routine occupations (mainly industrial blue-collar occupations), which exhibit decreasing wage premiums over time. However, my estimates indicate that premiums for males grew more rapidly than those for female workers within cognitive occupations and that these within-occupation differences in wage premiums growth largely dominate the compensating effect of gender differences in sorting across occupations.

We can conclude that when including part-time workers the gap in the changes in occupation premiums experienced by male and female workers is reduced, as the estimated premiums for female workers in interactive non-routine and manual non-routine occupations slightly increase with the presence of part-time workers. However, the results are quite similar for both samples, and the main conclusions remain unchanged.

Table B.1: Comparison of logarithms hourly wages approximated in SIAB administrative data with those on SOEP survey data

	Administrative data			SOEP survey data		
	Mean	Sd	N	Mean	Sd	N
Full-time						
Male	2.689	0.562	5,244,403	2.709	0.471	74,412
Females	2.379	0.551	2,505,647	2.471	0.487	30,992
Part-time						
Male	2.327	0.708	160,103	2.433	0.670	1,628
Females	2.435	0.523	1,314,490	2.371	0.520	23,381

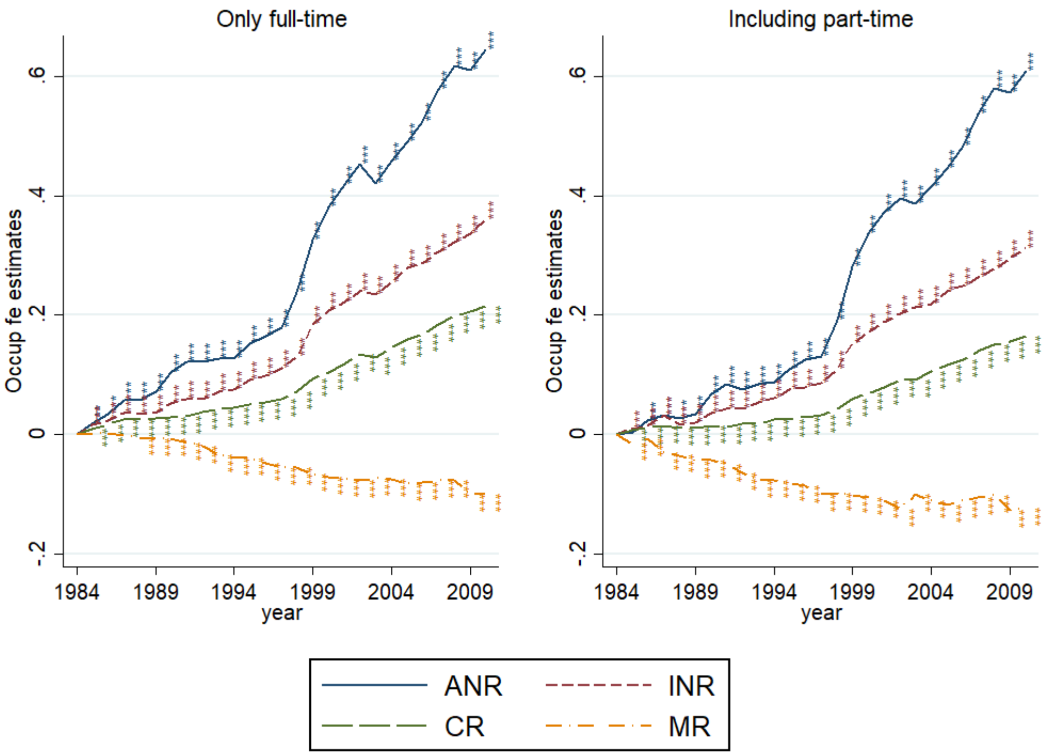
Notes: The logarithms of the hourly wages obtained in SOEP survey data are obtained through dividing the monthly labor income by the individual's declared weekly hours of work (multiplied by 4). The logarithm of hourly wages in SIAB data is obtained by dividing the administrative information of individual's daily wages by the information obtained in SOEP on the average hours worked by individuals of same sex and type of work (part-time/full-time) in each year (divided by seven to convert from hours per week to daily hours worked).

Table B.2: Decomposition of the changes in the gender wage premiums gap.

	Full-time only (1)	Full-time hourly wages (2)	Including part-time (3)
Gender gap in the base year	0.440	0.381	0.372
Change in gender gap in mean log wages	-0.082	-0.094	-0.105
A. Mixed by gender ocupation wage premiums			
$E[\theta_{jt}^{ALL} \| male]$	0.140	0.136	0.109
$E[\theta_{jt}^{ALL} \| female]$	0.237	0.236	0.199
$E[\theta_{jt}^{ALL} \| male] - E[\theta_{jt}^{ALL} \| female]$	-0.098	-0.100	-0.090
% of the initial wage gap	*-22.22*	*-26.29*	*-24.19*
B. Gender-specific occupation wage premiums			
$E[\theta_{jt}^{M} \| male]$	0.128	0.128	0.136
$E[\theta_{jt}^{F} \| female]$	-0.025	-0.012	0.032
$E[\theta_{jt}^{M} \| male] - E[\theta_{jt}^{F} \| female]$	0.152	0.140	0.104
% of the initial wage gap	*34.67*	*36.69*	*27.85*
C. Decomposition			
Sorting across occupations			
$E[\theta_{jt}^{F} \| male] - E[\theta_{jt}^{F} \| female]$	-0.024	-0.024	-0.023
% of the gender difference in occup premiums	*-16.06*	*-17.46*	*-22.59*
Within occupation differences			
$E[\theta_{jt}^{M} - \theta_{jt}^{F} \| male]$	0.177	0.164	0.127
% of the gender difference in occup premiums	*116.08*	*117.52*	*122.59*
Observations in last year	278,142	278,142	356,672

Notes: Column (1) includes only the sample of full-time workers and the estimated wage premiums are based on daily wages, Column (2) includes only full-time workers and uses the aproximate measure of hourly wages, and Column (3) includes both part-time and full-time workers and uses hourly wages. $E[\theta_{jt}^{ALL}|male]$ is the mean value of the mixed by gender occupation premia across male workers (considering male's distribution of occupations in the last year), while $E[\theta_{jt}^{ALL}|female]$ takes the mean across female workers. $E[\theta_{jt}^{M}|male]$ and $E[\theta_{jt}^{F}|female]$ are the mean values of male and female specific occupation premia across male and female workers respectively.

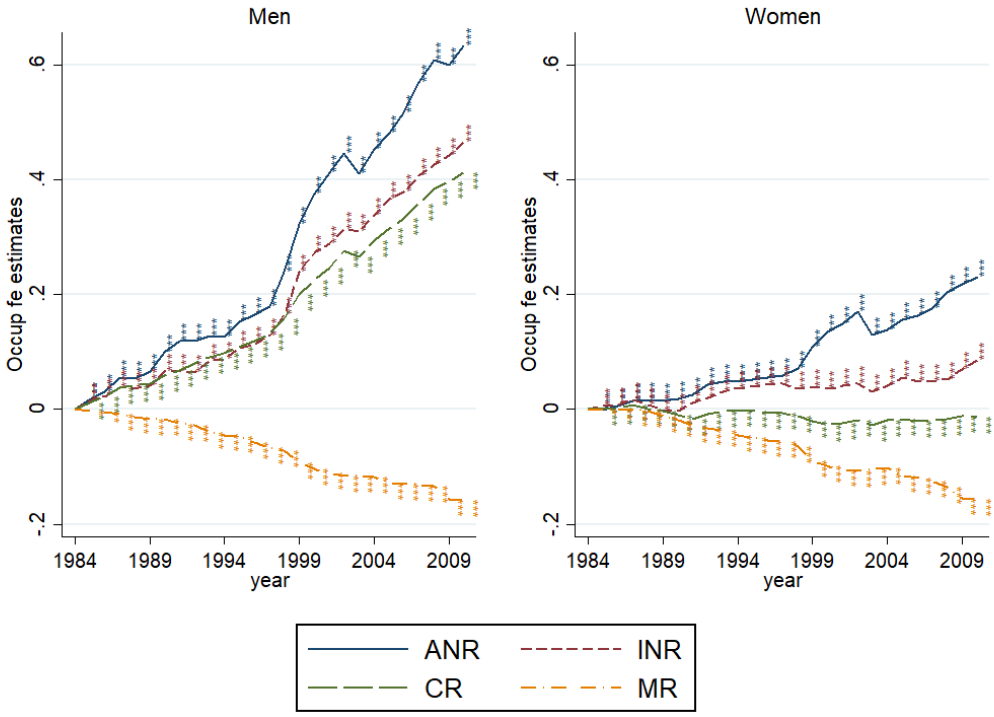
Figure B.1: Change in occupational premiums. Mixed-by-gender fixed effects.



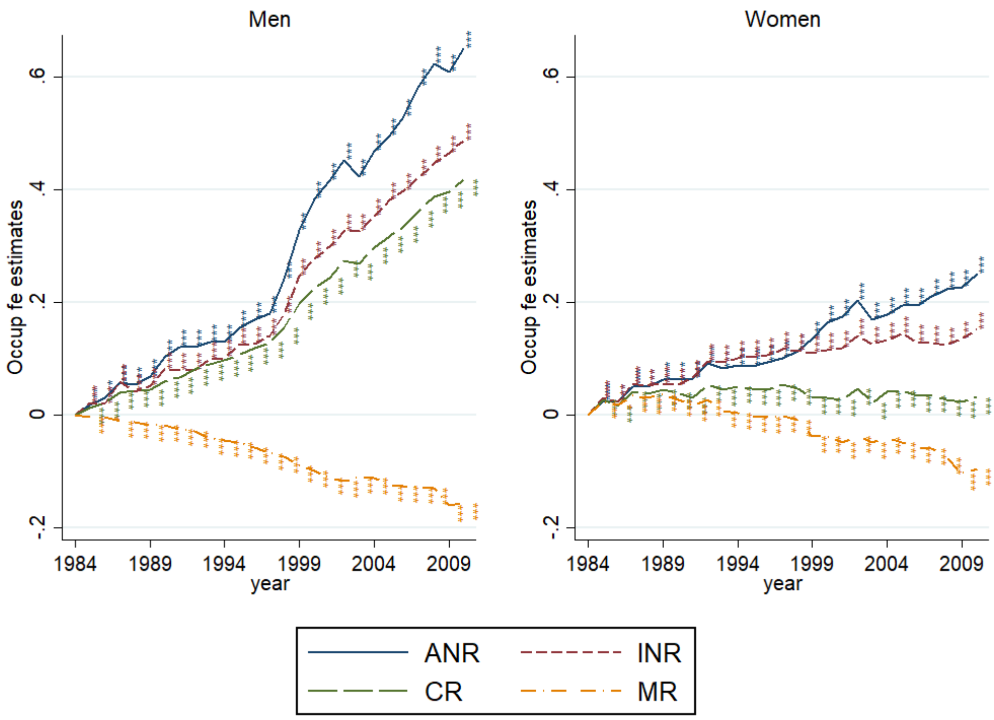
Notes: the figure plots the estimated coefficients on occupation-year dummies for a sample containing both male and female workers. The estimates for full-time workers are based on daily wage while the estimates for the sample including part-time workers uses the approximated measure for hourly wage. Stars denote the level at which the estimated coefficients are significantly different from zero (*** p<0.01, ** p<0.05, * p<0.1)

Figure B.2: Change in gender-specific occupational premiums

Panel A. Only full-time workers (using daily wage)



Panel B. Including part-time workers (using hourly wage)



Notes: the figure plots the estimated coefficients on occupation-year dummies for male and female workers. Stars denote the level at which the estimated coefficients are significantly different from zero (*** p<0.01, ** p<0.05, * p<0.1). They measure the change relative to 1984 in the occupation premia with respect to the analogous change of manual non-routine for men.

C Data preparation

C.1 Imputation of right censored wages

In the data coming from the Employee History (BeH), the variable *daily wage* shows the employee's gross daily wage, expressed in euros. It is calculated from the fixed-period wages reported by the employer and the duration of the (unsplit) original notification period in calendar days. Earnings exceeding the upper earnings limit for statutory pension insurance are only reported up to this limit. If the gross wage is higher than the current contribution limit, the amount of the ceiling is liable for the contribution and the wage information in this sample is censored at this limit.[29]

We assume that for individual i the wage in logs is given by:

$$y_i^* = x_i\beta + \varepsilon_i$$

$$where \;\; \varepsilon_i \overset{iid}{\sim} N(0,\sigma^2), \;\; i = 1, ..., n$$

As the wages in the SIAB are censored at the contribution limit a, we observe the wage $y_{obs,i} = y_i^*$ only if the wage is lower than the threshold a. If the wage has a greater value than a, it is censored, we observe the limit a instead of the true wage y_i^* :

$$y_i = \begin{cases} y_{obs,i} & if \;\; y_i^* \leq a \\ a & if \;\; y_i^* > a \end{cases}$$

This censoring affects 4.7% of the reported wages in the data. Especially for analyzing highly-skilled employees (with technical college degree or university degree) it is necessary to impute the missing wages.

Following Büttner and Rässler (2008) and Gartner and Rässler (2005) the censored data problem can be interpreted as a missing data problem. Let $Y = (Y_{obs}; Y_{mis})$ denote the random variables concerning the data with observed and missing parts. For all units with wages below the contribution limit each data record is complete, i.e., $Y = (Y_{obs}) = (X; wage)$. For every unit with a value equal to the contribution limit for its wage information we treat the data record as partly missing, i.e., $Y = (Y_{obs}; Y_{mis}) = (X; ?)$. X is observed for all units. Thus, we have to impute the missing data $Y_{mis} = wage$.

I follow the heteroscedastic single imputation approach described by Büttner and Rässler (2008). One of the advantages of the method they propose is that it does not presume homoscedasticity of the residuals. That is, they assume that the variation of income is smaller in lower wage categories than in higher categories, and therefore they propose an imputation approach considering heteroscedasticity.

In this approach the error variance is related to a number of exogenous variables, gathered in

[29] For reference, in 2008 the contribution limit in the unemployment insurance system is fixed in Western Germany at a monthly income of 5,300 euros.

a vector w. They suggest to use a GLS model for truncated variables (ex: *intereg* in STATA) to estimate the parameters of the imputation model, β, and furthermore γ, describing the functional form of the heteroscedasticity. Then the imputation can be done by draws from a truncated normal distribution.

$$z_i \sim N_{trunc_a}(x_i\hat{\beta}, \hat{\sigma_i^2})$$

$$where\ \hat{\sigma_i^2} = e^{w_i}\hat{\gamma}\ \ if\ \ y_i = a\ \ for\ \ i = 1, ..., n$$

where w is a vector of observed variables that is a subset of x.

In vector x I include six age categories, three skill categories, interactions of skill and age categories, place of work (federal level), sector (17 categories, based on Industry classification 1993) and a dummy for part-time. Vector w contains the age and skill categories.[30] I applied this imputation separately for males and females and for each year (1975-2010). To consider the heteroscedastic structure of the residuals, they propose to draw a random variable for the individual variances for every person.

C.2 Education imputation

The reporting of the educational degree of an employee to the social insurance agencies by the employer is part of the compulsory notification system. However, the educational degree has no consequences concerning obligations or claims out of the social security neither for the employer nor for the employee. Therefore, this variable can be regarded as less reliable than other variables like wages. More specifically the education variable in the SIAB data has two main problems: missing and inconsistent education information.

To deal with the problem of missing values and/or inconsistent sequences of educational reports, I follow Fitzenberger et al. (2005), who propose to use deductive imputation methods that make use of the panel structure of the data set. In addition, information contained in the variables employment status and age is considered when constructing the corrected education variables.

They develop three different imputation methods based on different assumptions on nature of the reporting process. I follow their Imputation procedure 1 (IP1), which assumes that underreports are the only possible source of inconsistencies as some employers (for reasons that are not observable to the researcher) do not report the actual educational degree of the employee but the degree required for the position. For a given person IP1 extrapolates every degree which is reported for the first time and is higher than the degrees reported previously, to subsequent spells with lower or missing education information.

This extrapolation procedure is based on three hypothesis: 1) after having attained an educational degree, individuals keep their degree; 2) the educational degree remains in general almost

[30] Gartner and Rässler (2005) include potential experience (linear, quadratic and cubic), 6 educational levels, 11 occupational groups according to Blossfeld (1985), 12 categories of firm size, 15 industrial categories. The imputation is done separately for males and females and for each year. Büttner and Rässler (2008) include 6 educational levels, age and age square, nation

constant once a person has entered working life; and 3) employers have to report the highest attained degree.

As long as one cannot completely rule out the possibility of overreports, this procedure is likely to induce an upward bias in the corrected education variable. Therefore, they develop two further imputation procedures that distinguish between reliable reports that are used for extrapolation and unreliable reports that have to be discarded. They use heuristic rules to identify valid education information in a conservative way.[31] These other alternatives for imputation tend to under-report education. In their paper, Fitzenberger et al. (2005) compare in three typical labor economic applications the data resulting from the three imputation procedures to the original data, and they find some evidence in favor of the hypothesis that underreporting of educational degrees is a more severe problem than overreporting. Their evidence suggests that, in fact, employers tend to report the degree required for the position rather than the highest qualification attained by the employee.

C.3 Task approach classification of occupations

The classification of occupations that I use in this paper follows the classification developed by Spitz-Oener (2006). To operationalize the task categories, I use data on tasks coming from the Qualification and Career Survey, which is a survey of employees carried out by the German Federal Institute for Vocational Training (BIBB). It includes five cross sections launched in 1979, 1985/86, 1991/92, 1998/99 and 2006, each covering a representative sample of about 30,000 workers (men and women). I restict my sample to men employed in West Germany who are aged between 18 and 65, and of German nationality to make the sample consistent across waves.

While the 1979 wave covers approximately 90 activities, the number of activities decreased to 19 in the 2012 wave. In order to create a task intensity measure that is consistent over time, I followed previous work which merged some of the activities into one variable in order to deal with the changing definitions of the variables and to maintain a total number of activities which is similar in each survey. I arrived to 17 longitudinally consistent tasks, and classified them into the five dimensions proposed by Spitz-Oener (2006)(see Table C.1).

[31] Imputation procedure 2 assumes that the observed frequency of a reported degree can be interpreted as a sign of its reliability. Imputation procedure 3, on the contrary, is based on the assumption that reporting errors are serially correlated. In this case, a change of the reported educational degree may reveal some information on the reliability of the employer issuing the reports. Imputation procedure 1 possibly overstates education through a ratched effect when overreporting occurs.

Table C.1: Classification of tasks

Category	Tasks
Analytical non-routine	Developing, researching, constructing, designing. Collecting information, investigating, documenting. Programming a Computer. Apply legal knowledge.
Interactive non-routine	Buying, selling, purchasing, advertising, public relations, promoting. Training, teaching, providing advice and information. Organizing, deciding, coordinating, planning, managing, hiring and controlling employees, negotiating.
Cognitive routine	Measuring, testing, quality control. Calculating, accounting. Typing, forms.
Manual routine	Producing goods. Transporting, storing, sorting, shipping, packing, loading, delivering, controlling vehicles. Monitoring, operating and controlling machines.
Manual non-routine	Repair or refubrish machinery, houses or vehicles. Renovate or restore. Entertaining, accommodating, preparing food. Nursing, caring, healing. Cleaning, removing waste, ironing, washing. Protecting, guarding, controlling traffic, security.

Notes: The tasks are taken from the BIBB surveys and the classification follows Spitz-Oener (2006).

To construct a single measure of the different tasks performed by individuals it is necessary to calculate an index of task intensity. In previous work, two different task intensity measures were developed using these data. Spitz-Oener (2006) proposed an index of "task intensity", which determines the degree to which a single task dimension is necessary, to perform a specific occupational activity when compared to another occupational activity:

$$SO_{ijt} = \frac{N^{\underline{o}}\, activities\, in\, category\, j\, performed\, by\, i\, at\, t}{Total\, N^{\underline{o}}\, of\, activities\, in\, j\, at\, t}$$

where j defines de borad occupational categories, $j = \{ANR, INR, CR, MR, MNR\}$ and t is the year of the survey (1979, 1985/86, 1991/92, 1998/99 and 2012).

Antonczyk et al. (2009) developed an index of "task composition", which specifies the shares of the different tasks in an individual's occupational activity:

$$AFL_{ijt} = \frac{N^{\underline{o}}\, activities\, in\, category\, j\, performed\, by\, i\, at\, t}{Total\, N^{\underline{o}}\, of\, activities\, performed\, by\, i\, at\, t}$$

For example, SO task index would tell us that the job of a manager requires relatively high analytical skills compared to the activities of a cleaner. The AFL index, would reflect the fact that in addition to high analytical demands, a manager must also meet high requirements in the other task dimensions. Therefore, the share of analytical activities is smaller compared with a researcher, for instance, who must perform analytical activities nearly exclusively. Both indices are extremely sensitive to the number of characteristics included in the survey. I calculate both indices and take means of each category at the occupational level (3 digits of Kldb 1988) in order to be able to classify occupations in the SIAB data. The final classification is based on the highest value of each index at the occupational level.

Tables C.5 to C.8 in the Appendix contain the detailed mapping of occupations into the task-based occupation categories.

Table C.2: Most relevant occupations in terms of employees in each category. Men

	Analytical non-routine	N	% in the group
628	Other technicians	137,037	14.46
774	Data processing specialists	134,584	14.20
602	Electrical engineers	79,557	8.40
622	Electrical engineering technicians	75,236	7.94
601	Mechanical, motor engineers	74,597	7.87
	Interactive non-routine	N	% in the group
751	Entrepreneurs, managing directors	152,297	29.46
629	Foremen, master mechanicals	70,638	13.67
841	Physicians	38,371	7.42
752	Managements consultants, organisers	26,806	5.19
862	Home wardens, social work teachers	24,062	4.65
	Cognitive routine	N	% in the group
781	Office specialists	438,627	50.30
691	Bank specialists	123,693	14.18
521	Good examiners, sorters, n.e.c.	53,372	6.12
694	Life, property insurance specialists	52,240	5.99
772	Accountants	32,673	3.75
	Manual routine	N	% in the group
714	Motor vehicle drivers	360,232	9.88
311	Electrical fitters, mechanics	187,752	5.15
744	Stores, transport workers	156,600	4.30
441	Bricklayers	137,288	3.77
273	Engine fitters	125,710	3.45
	Manual non-routine	N	% in the group
682	Salespersons	118,284	12.95
681	Wholesale and retail trade buyers	113,215	12.39
687	Commercial agents, travelers	92,480	10.12
511	Painters, lacquerers, construction	72,338	7.92
51	Gardeners, garden workers	61,664	6.75

Notes: The table presents the first five occupations in each group that concentrate the highest number of employees. The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).

Table C.3: Most relevant occupations in terms of employees in each category. Women

	Analytical non-routine	N	% in the group
774	Data processing specialists	5,877	18.31
635	Technical draughtsperson	4,628	14.42
628	Other technicians	2,760	8.6
882	Economic and social scientists	2,329	7.25
821	Journalists	1,606	5
	Interactive non-routine	N	% in the group
864	Nursery teachers, child nurses	17,401	20.56
861	Social workers, care workers	14,769	17.45
862	Home wardens, social work teachers	9,315	11.01
841	Physicians	4,022	4.75
751	Entrepreneurs, managing directors	3,739	4.42
	Cognitive routine	N	% in the group
781	Office specialists	131,430	54.09
856	Medical receptionists	23,142	9.52
691	Bank specialists	18,196	7.49
782	Steneographers, short-hand typists	17,658	7.27
784	Office auxiliary workers	6,472	2.66
	Manual routine	N	% in the group
744	Stores, transport workers	6,126	8.49
321	Electrical appliance	5,404	7.49
732	Postal deliverers	3,852	5.34
323	Metal workers	3,779	5.24
322	Other assemblers	3,459	4.8
	Manual non-routine	N	% in the group
682	Salespersons	52,898	29.05
853	Nurses, midwives	31,699	17.41
933	Household cleaners	25,746	14.14
411	Cooks	11,183	6.14
681	Wholesale and retail trade buyers	10,634	5.84

Notes: The table presents the first five occupations in each group that concentrate the highest number of employees. The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).

Table C.4: Proportion of females in each occupational group

	N of observations	Percentage of women
Analytical NR	1,160,402	16.5
Interactive NR	1,105,396	49.6
Cognitive R	2,766,367	67.6
Manual R	4,365,876	15.5
Manual NR	2,407,896	60.3

Table C.5: Occupations classified as Cognitive non-routine

Analytical non-routine

Code	*Description*
283	Aircraft mechanics
284	Precision mechanics
303	Dental technicians
601	Mechanical, motor engineers
602	Electrical engineers
603	Architects, civil engineers
604	Survey engineers
605	Mining, metallurgy, foundry engineers
606	Other manufacturing engineers
607	Other engineers
611	Chemists, chemical engineers
612	Physicists, physics engineers, mathematicians
621	Mechanical engineering technicians
622	Electrical engineering technicians
623	Building technicians
624	Measurement technicians
625	Mining, metallurgy, foundry technicians
626	Chemistry, physics technicians

Code	*Description*
627	Remaining manufacturing technicians
628	Other technicians
631	Biological specialists
632	Physical and mathematical specialists
633	Chemical labouratory assistants
635	Technical draughtsperson
721	Navigating ships officers
722	Technical ships officers, ships engineers
726	Air transport occupations
774	Data processing specialists
811	Arbitrators
812	Judicial administrators
814	Judicial enforcers
821	Journalists
836	Interior, exhibition designers, window dressers
881	Economic and social scientists, statisticians
882	Humanities specialists, n.e.c.
883	Scientists n.e.c.

Interactive non-routine

Code	*Description*
31	Managers in agriculture and animal breeding
32	Agricultural engineers, agriculture advisors
52	Garden architects, garden managers
304	Ophthalmic opticians
629	Foremen and other operations managers
692	Building society specialists
693	Health insurance specialists (not social security)
703	Publicity occupations
704	Brokers, property managers
705	Landlords, agents, auctioneers
751	Entrepreneurs, managing directors, divisional managers
752	Management consultants, organisers
761	Members of Parliament, Ministers, elected officials
762	Senior government officials
813	Legal representatives, advisors
831	Musicians
832	Artists' agents
833	Visual, commercial artists

Code	*Description*
844	Pharmacists
851	Non-medical practitioners
852	Masseurs, physiotherapists and related occupations
855	Dietary assistants, pharmaceutical assistants
861	Social workers, care workers
862	Home wardens, social work teachers
863	Work, vocational advisers
864	Nursery teachers, child nurses
871	University teachers, lecturers at higher technical schools and academies
872	Gymnasium teachers
873	Primary, secondary (basic), special school teachers
874	Technical, vocational, factory instructors
875	Music teachers, n.e.c
876	Sports teachers
877	Other teachers
892	Nuns, friars and other religious associate professionals
891	Ministers of religion
911	Restaurant, inn, bar keepers, hotel

Table C.6: Occupations classified as Cognitive routine

Code	*Description*	*Code*	*Description*
142	Chemical laboratory workers	753	Chartered accountants, tax advisers
423	Other beverage makers, tasters	763	Association leaders, officials
521	Goods examiners, sorters, n.e.c.	771	Cost accountants, valuers
683	Publishing house dealers, booksellers	772	Accountants
691	Bank specialists	773	Cashiers
694	Life, property insurance specialists	781	Office specialists
701	Forwarding business dealers	782	Stenographers, shorthand-typists, typists
702	Travel agency clerks, attendants, organizers and guides	783	Data typists
706	Cash collectors, cashiers, ticket sellers, inspectors	784	Office auxiliary workers
712	Railway controllers, conductors	803	Safety testers
713	Other traffic controllers, conductors	805	Health-protecting occupations
731	Post masters	822	Interpreters, translators
733	Radio operators	823	Librarians, archivists, museum specialists
734	Telephonists	856	Medical receptionists
		857	Medical labouratory assistants

Notes: The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).

Table C.7: Occupations classified as Manual routine

Code	*Description*
11	Farmers
12	Winegrowers
22	Fishermen
41	Land workers
42	Milkers
62	Forest workers, forest cultivators
71	Miners
72	Mechanical, electrical, face workers, shot firers
81	Stone crushers
82	Earth, gravel, sand quarriers
83	Oil, natural gas quarriers
91	Mineral preparers, mineral burners
101	Stone preparers
102	Jewel preparers
111	Stoneware, earthenware makers
112	Shaped brick, concrete block makers
121	Ceramics workers
131	Frit makers
132	Hollow glassware makers
133	Flat glass makers
134	Glass blowers (lamps)
135	Glass processors, glass finishers
141	Chemical plant operatives
143	Rubber makers, processors
144	Vulcanisers
151	Plastics processors
161	Paper, cellulose makers
162	Packaging makers
163	Book binding occupations
164	Other paper products makers
171	Type setters, compositors
172	Printed goods makers
173	Printers (letterpress)
174	Printers (flat, gravure)
175	Special printers, screeners
176	Copiers
177	Printer's assistants
181	Wood preparers
182	Wood moulders and related occupations
183	Wood products makers
184	Basket and wicker products makers
191	Iron, metal producers, melters
192	Rollers
193	Metal drawers
201	Moulders, coremakers
202	Mould casters
203	Semi-finished product fettlers and other mould casting occupations
211	Sheet metal pressers, drawers, stampers
212	Wire moulders, processors
213	Other metal moulders (non-cutting deformation)
221	Turners
222	Drillers
223	Planers
224	Borers
225	Metal grinders
226	Other metal-cutting occupations
231	Metal polishers
232	Engravers, chasers
233	Metal finishers
235	Enamellers, zinc platers and other metal surface finishers
241	Welders, oxy-acetylene cutters
242	Solderers
244	Metal bonders and other metal connectors
251	Steel smiths
252	Container builders, coppersmiths and related occupations
261	Sheet metal workers
262	Plumbers
263	Pipe, tubing fitters
270	Locksmiths, not specified
271	Building fitters
272	Sheet metal, plastics fitters
273	Engine fitters
274	Plant fitters, maintenance fitters
275	Steel structure fitters, metal shipbuilders
281	Motor vehicle repairers
282	Agricultural machinery repairers
285	Other mechanics
286	Watch-, clockmakers
291	Toolmakers
301	Precision fitters n.e.c
302	Precious metal smiths
311	Electrical fitters, mechanics
312	Telecommunications mechanics, craftsmen

Notes: The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).

Code	Description
313	Electric motor, transformer fitters
315	Radio, sound equipment mechanics
321	Electrical appliance, electrical parts assemblers
322	Other assemblers
323	Metal workers (no further specification)
331	Spinners, fibre preparers
332	Spoolers, twisters, ropemakers
341	Weaving preparers
342	Weavers
343	Tufted goods makers
344	Machined goods makers
345	Felt makers, hat body makers
346	Textile processing operatives braiders
352	Clothing sewers
353	Laundry cutters, sewers
355	Hat, cap makers
366	Sewers, n.e.c.
357	Other textile processing operatives
361	Textile dyers
362	Textile finishers
371	Leather makers, catgut string makers
373	Footwear makers
374	Coarse leather goods finishers, truss makers
378	Skin processing operatives
391	Bakery goods makers
392	Confectioners (pastry)
401	Butchers
402	Meat, sausage goods makers
403	Fish processing operatives
412	Ready-to-serve meals, fruit, vegetable preservers, preparers
421	Wine coopers
422	Brewers, maltsters
424	Tobacco goods makers
431	Milk, fat processing operatives
432	Flour, food processors
433	Sugar, sweets, ice-cream makers
441	Bricklayers
442	Concrete workers
451	Carpenters
452	Roofers
453	Scaffolders
461	Paviors
462	Road makers

Code	Description
463	Tracklayers
464	Explosives men (except shotfirers)
465	Land improvement, hydraulic engineering workers
466	Other civil engineering workers
470	Building labourer, general
471	Earth movers
472	Other building labourers, building assistants, n.e.c.
481	Stucco workers, plasterers, rough casters
486	Screed, terrazzo layers
492	Upholsterers, mattress makers
501	Carpenters
502	Model, form carpenters
503	Cartwrights, wheelwrights, coopers
504	Other wood and sports equipment makers
512	Goods painters, lacquerers
513	Wood surface finishers, veneerers
514	Ceramics, glass painters
522	Packagers, goods receivers, despatchers
531	Assistants (no further specification)
541	Generator machinists
542	Winding engine drivers, aerial ropeway machinists
543	Other machinists
544	Crane drivers
545	Earthmoving plant drivers
546	Construction machine attendants
547	Machine attendants, machinists' helpers
548	Stokers
549	Machine setters (no further specification)
634	Photo labouratory assistants
711	Railway engine drivers
714	Motor vehicle drivers
723	Deck seamen
724	Inland boatmen
725	Other water transport occupations
732	Postal deliverers
741	Warehouse managers, warehousemen
742	Transportation equipment drivers
743	Stowers, furniture packers
744	Stores, transport workers
834	Scenery, sign painters

Notes: The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).

Table C.8: Occupations classified as Manual non-routine

Code	*Description*	*Code*	*Description*
21	Animal breeders	687	Commercial agents, travellers
43	Family-member land workers, n.e.c.	688	Mobile traders
44	Animal keepers and related occupations	715	Coachmen
51	Gardeners, garden workers	716	Street attendants
53	Florists	791	Factory guards, detectives
243	Riveters	792	Watchmen, custodians
305	Musical instrument makers	793	Doormen, caretakers
306	Doll makers, model makers, taxidermists	794	Domestic and non-domestic servants
314	Electrical appliance fitters	801	Soldiers, border guards, police officers
351	Cutters	802	Firefighters
354	Embroiderers	804	Chimney sweeps
372	Shoemakers	838	Performers, professional sportsmen, auxiliary artistic occupations
375	Fine leather goods makers	853	Nurses, midwives
376	Leather clothing makers and other leather processing operatives	854	Nursing assistants
377	Hand shoemakers	893	Religious care helpers
411	Cooks	901	Hairdressers
482	Insulators, proofers	902	Other body care occupations
483	Tile setters	912	Waiters, stewards
484	Furnace setter, air heating installers	913	Others attending on guests
485	Glaziers	923	Other housekeeping attendants
491	Room equippers	931	Laundry workers, pressers
511	Painters, lacquerers (construction)	932	Textile cleaners, dyers and dry cleaners
681	Wholesale and retail trade buyers, buyers	933	Household cleaners
682	Salespersons	934	Glass, buildings cleaners
684	Druggists /chemists (pharmacy)	935	Street cleaners, refuse disposers
685	Pharmacy aids	936	Vehicle cleaners, services
686	Service-station attendants	937	Machinery, container cleaners and related occupations

Notes: The occupational codes correspond to the German classification of occupations (KldB) of 1988 (*Klassifizierung der Berufe* 1988).