DOES THE GENDER COMPOSITION OF AN OCCUPATION AFFECT WAGES?

Nicola Fuchs-Schündeln*, Franziska Riepl**, and Alexandra Spitz-Oener***

August 2025

Abstract

We document that increased occupational female shares causally lead to lower occupational wages. When the Berlin Wall fell, women in the German Democratic Republic were not only more likely to participate in the labor market than their West German counterparts, but were also distributed differently across occupations. Exploiting German reunification as a natural experiment, we use variation in the gender composition of East and West German occupations as an instrument for changes in occupational female shares in West Germany. We show that the gender composition of an occupation affects its wages. The adverse effects of increasing female shares are pervasive and all-encompassing, and not driven by changes in skill requirements or the task content in occupations. Overall, the results suggest that decreasing occupational segregation might be effective in closing the gender wage gap. However, the mechanism is quite different than the one that is typically discussed.

Acknowledgements: We thank Martin Groiss, Violetta Kuzmova, and Arnim Seidlitz for outstanding research assistance. We thank Josh Angrist, David Autor, Sandy Black, Leah Boustan, Ellora Derenoncourt, Christian Dustmann, Claudia Goldin, Claudia Olivetti, Barbara Petrongolo, Benjamin Schoefer, and participants at the NBER Summer Institute Gender in the Economy as well as at multiple seminar presentations for excellent suggestions. Fuchs-Schündeln gratefully acknowledges financial support by the ERC through Consolidator Grant No. 815378 and by the DFG through the Gottfried Wilhelm Leibniz-Preis. Spitz-Oener gratefully acknowledges financial support by the DFG through CRC TRR 190 Project No. 280092119, and by the Rockwool Foundation Berlin (RFBerlin), Project No. 2013.

*WZB Berlin Social Science Center, Goethe University Frankfurt, and CEPR.
**WZB Berlin Social Science Center.
***Humboldt-Universität zu Berlin, RFBerlin, and Institute for Employment Research (IAB).

1 Introduction

Despite striking long-run improvements, evidence from recent decades suggests that women's progress in closing the gender pay gap has considerably slowed, if not halted, in industrialized countries. Gender differences in occupational choices are one important driver of the persistent gender pay gap.[1] Women continue to be overrepresented in low-paying occupations and men in high-paying occupations. To further reduce the gender wage gap, it thus seems crucial for men and women to be more equally distributed across occupations. Consistent with this widespread assessment, policies that aim at increasing female labor force participation *per se* have been augmented by additional policies that foster female labor force participation in specific occupational fields, the prime example being the STEM fields – Science, Technology, Engineering and Mathematics.

Against this background, it seems crucial to understand how changes in the distribution of men and women across occupations, i.e., changes in the female share of occupations, affect wages of men and women in the respective occupations. Previous research documents a robust negative correlation between the female share of occupations and relative wages for both male and female workers in that occupation in the context of different countries and periods (see e.g. Levanon et al., 2009, Macpherson and Hirsch, 1995, and Murphy and Oesch, 2016). Based on this finding, the sociological literature has developed the 'devaluation hypothesis' (see e.g. England et al., 1988). This hypothesis states that society values jobs carried out by women less than jobs carried out by men; lower prestige is associated with lower wages, and thus, an increase in the female share of an occupation leads to lower wages for all workers in that occupation. However, to date, we do not know in which direction causality runs in the negative correlation. The complexity of the phenomenon creates threats for causal identification. On the one hand, men and women change their occupations over time for various supply-side reasons, including preferences, educational attainment, social norms, and policies. On the other hand, employers also change the gendered occupational demand for labor services due to changing occupational task requirements, among other things. To give a concrete example, some occupations might increase their amenities over time in combination with a compensating negative wage differential; if these amenities, e.g., job flexibility, are especially attractive to women, this could increase the female share in these occupations (see, e.g., Felfe, 2012). Therefore, one needs a shock to female (vs. male) labor supply to identify a causal effect of the female share on wages.

[1] Blau and Kahn (2017) show that the full-time workers' female-to-male earnings ratio increased from around 60 percent in the United States in the 1980s to around 75 percent at the beginning of the 2000s, with a much slower increase after that. They assess the relative importance of different explanations of the gender pay gap. Bertrand (2020) provides international evidence on the gender wage gap and illustrates the importance of differences in occupational choices between men and women. Hsieh et al. (2019) assess the loss in aggregate productivity associated with an unequal gender or race distribution across occupations.

This paper uses such a shock for causal identification by exploiting German reunification as a gendered labor supply shock to the West German labor market. Specifically, we instrument the change in the actual female share in an occupation in West Germany between the 1980s and the 1990s with the change in the female share induced by the opening of the West German labor market to the East German workforce. The fall of the Berlin Wall on November 9, 1989, the subsequent collapse of the GDR, and the merger of the East and West German labor markets came unexpectedly for East German workers. The pre-1989 training occupations of East German men and women were determined by the education and training system in the GDR and various GDR-specific policies. They were not chosen according to expectations concerning labor market prospects in West Germany after reunification or according to female shares in occupations in West Germany. We document that occupational female shares differed substantially between East and West Germany in the 1980s. We argue that this is due to a combination of factors, like differences in policies, norms, and industrial structures, between the two countries.

Reunification then acted as a gendered labor supply shock and added the pool of East German men and women as potential workers in the West German labor market. Our instrumental variable is an adjusted version of the instrument used by Prantl and Spitz-Oener (2020). We instrument the change in the female share on the occupation-age level pre- and post-reunification in the West with the change in the gendered composition of the occupation-age cell induced by the pool of East Germans in the age group and the training occupations related to the respective occupation-age cell. These are the potential additional workers added to the occupation-age cell due to reunification. Thereby, we can analyze the causal impact of a change in the occupational female shares on the relative wages of West German workers in the respective occupation. Importantly for identification, the setting allows us to isolate the causal effect of an exogenous change in the gender composition of occupations on wages from the impact of changes in labor demand or endogenously arising changes in labor supply on wages. When doing this, we also consider the effect of the general labor supply shock on wages in the context of German reunification and specifically isolate the impact of changes in occupational female shares.

Figure 1 provides suggestive evidence of the effect on wages of the gendered occupational shock we exploit. The upper panel shows the normalized average gross wages for full-time employed men in West Germany from 1985 to 1999 for two occupational groups.[2] The first group comprises the 40 percent of occupations for which East Germans' female share in training occupations exceeded the female share in current occupations in the West in the 1980s the most (Top 40%, red line). The second group comprises the opposite 40 percent of occupations, for which the female share difference between East and

[2]This graph relies on the raw data and thus does not include any controls (e.g., the set of fixed effects we later add to the regression analysis).

West was the smallest (Bottom 40%, blue line). As the figure shows, wages in both groups of occupations moved in lock-step before 1989. However, after the fall of the Berlin Wall in November 1989, the wages of West Germans in occupations with the largest East-West female share difference declined relative to those in occupations with the smallest difference. Thus, the raw data suggests that the changes in occupational female shares caused by German reunification might have affected the occupational wage structure. While the raw data indicates that these effects died out in the late 1990s, in the regressions we find that not to be the case.

We compare the evolution of normalized male wages of the same occupations using U.S. data in the lower panel.[3] While the U.S. data are somewhat noisy, for the U.S., wages in the two groups of occupations evolved very differently from those in West Germany throughout the observation period. The difference in observed patterns for the two countries suggests that the groups of occupations that were differentially hit by the gender shock owing to the merger of the two German labor markets after 1989 do not incidentally reflect occupations that were subject to other forces, such as technological changes, that affected occupations differentially during this time (see Böhm et al., 2024, and the comprehensive list of references cited there).

Our research design furthermore allows us to rule out other explanations, such as changes in labor demand or changes in amenities that attract more women and are related to lower wages due to compensating wage differentials. This is done by instrumenting the change in the female share, but is also facilitated by including several fixed effects. We can also rule out that the increased supply of workers in the context of German reunification drives the negative impact on wages.

The instrumental variable (IV) results suggest that a higher female share is indeed causally related to lower wages. Specifically, the results suggest that a one percentage point increase in the female share of an occupation-age cell leads to 0.7 percent lower wages. These quantitative effects are large enough to caution that policies aimed at increasing female participation in male-dominated occupations may still be effective in increasing female wages, but face a counteracting force.

We show that the effects are pervasive. For example, they do not depend on the initial female share in the occupation, and are equally large in manufacturing and services. We also find little heterogeneity based on monopsony power in the local labor market. Despite this general pervasiveness of the effect, the results also suggest smaller adverse effects for worker groups who are more likely to be subject to formalized wage-setting processes either because they work at specific types of establishments (e.g., large establishments or high AKM fixed-effect establishments) or because labor law makes their wages inflexible (e.g., older workers who typically have long-term work contracts). We

[3] We use a crosswalk from German to US occupations to group the occupations identically for the two countries. More details on the construction of this figure can be found in Appendix A.

also find smaller negative effects in competitive markets relative to markets in which firm entry is restricted. With few exceptions, however, the quantitative importance of these heterogeneities is small relative to the baseline negative effect.

Strikingly, the negative wage effects are nearly as large – about 95% – in districts where the female share in the occupation did NOT increase as in those where it actually did increase. This strongly suggests that the mechanism at play is not driven by actual changes in job content, but rather by shifting perceptions about the skill requirements of occupations as they become increasingly female. The former is corroborated by analyses of the task content of occupations. We find either no effect of the changing female share in an occupation on the tasks performed by workers in the occupation, or effects that are quantitatively too small to explain the changes in wages. In addition, the explanation of shifting perceptions is corroborated by analyses of prestige score data, which suggest that the prestige of occupations declines as they become more female. Together, these findings point to a generalized devaluation process triggered by the feminization of occupations, likely rooted in status-based beliefs about gender and skill.

We find one important heterogeneity when it comes to the question of how decreases in occupational segregation can impact the gender wage gap. Specifically, we find that the adverse wage effect of increases in the occupational female shares is more pronounced for men than for women. For men, a one percentage point increase in the occupation-age cell's female share leads to 0.8 percent lower wages. The coefficient for female wages is also negative but smaller, at 0.48 percent. Evaluated at the mean change of the female share between 1985/86 and 1998/99, male wages decreased by about 2.33 percent, and female wages increased by about 0.24 percent owing to the changing female shares.[4] This result suggests an unintended side effect of policies intended to reduce the gender wage gap by reducing occupational segregation. These policies are typically discussed in the context of women catching up to men's higher wages by moving into higher-paying, male-dominated occupations. Our results suggest that the reduction in men's wages resulting from the increase in the female share in occupations also reduces the gender wage gap. However, the mechanism is more nuanced and quite different from what is typically put forward.

This latter, unintended effect is important as it might affect how willing men are to welcome women as colleagues. Again, it is consistent with the notion that occupational prestige is negatively affected when women enter an occupation, potentially leading to men's reluctance to have women enter predominantly male occupations (as put forward, for example, in the pollution theory by Goldin, 2014).

The findings are consistent with recent evidence suggesting that men assess the workplace quality as deteriorating when women enter the occupation, even though "hard"

[4]Occupational female shares of men increased by 2.91 percentage points, and women's declined by 0.49 percentage points between 1985/86 and 1998/99.

measures would not suggest so. Specifically, Greenberg et al. (2024) investigate how the entry of women in a previously mostly male-only occupational domain (combat units of the U.S. military) affects the assessment of male workers in those occupations. The authors show that including women in the previously male-only combat units did not affect men's job performance or behavioral outcomes such as retentions, promotions, or criminal charges. Despite the lack of effects for these "well-measured" outcomes from personal records, survey results indicate a negative impact on the male soldiers' perceptions of workplace quality. Unlike our setting, wages in that environment are set and cannot be adjusted.

More generally, our study contributes to an area of the literature that, so far, has struggled with moving toward establishing a causal effect of occupational female shares on wages. Levanon et al. (2009) rely on Granger causality and find that the lagged female share correlates with wages, but not lagged wages with the female share. Macpherson and Hirsch (1995) document a negative correlation between the female share in occupations and male and female wages for the US in CPS data. They show that the strength of the correlation is reduced, even though it remains negative and significant, once individual fixed effects are included to control for unobserved time-constant individual heterogeneity, a result that we replicate. This indicates that part of the negative correlation might be due to sorting.[5] Murphy and Oesch (2016) provide similar evidence based on individual fixed effects regressions for Britain, Germany, and Switzerland.

A recent paper by Harris (2022) relies on US census data over 5 decades and uses a shift-share instrument to establish causality. The instrument keeps the proportion of workers from each education level in an occupation and the likelihood of men and women of a certain education level to work in a particular occupation fixed at the 1980 level (the share part). The shift part of the instrument consists of the changes in the gender composition for each education group, specifically the relative increase in female higher education from 1960 to 2010. He adds several control variables to address the remaining concerns of endogeneity. He finds that a 10 percentage point increase in the female share of an occupation leads to a 7 to 8 percent decrease in female and male wages. In contrast to a shift-share instrument, we use a more explicit gendered labor supply shock stemming from a quasi-experimental setting. In addition, our setup allows us to explore various mechanisms potentially driving the results.

The rest of the paper is structured as follows. Section 2 provides the institutional background and describes differences in the gendered occupational choices in East and West Germany. Section 3 introduces our two data sources, defines the main variables, and describes the sample selection. Section 4 then describes the empirical strategy, including the construction of the instruments, and presents the main results. Section 5 shows the

[5] Discrimination in the past or present might have resulted in lower wages for "female" occupations, and consequently, these occupations attract men and women with lower unobserved skills.

robustness of the results and explores potential mechanism, before Section 6 concludes.

2 Background

The German Democratic Republic (referred to as East Germany here, or "the GDR") and the Federal Republic of Germany (West Germany, or "the FRG") were officially unified on October 3, 1990, after four decades of division from 1949 to 1989. The GDR was a state socialist society and a centrally planned economy, and the FRG was a social-market economy. Thus, the political and economic differences between the two parts during the separation were very pronounced.

Employment and family policies were also important areas where the two parts of Germany differed, producing stark differences in the gendered nature of work (e.g., Rosenfeld et al., 2004). While West Germany supported the male-breadwinner model by introducing a joint taxation system for married couples in 1958, among other measures, and provided limited childcare opportunities, women in the GDR were not only actively encouraged to work; they had the duty to work.[6] Consistent with the latter, universal childcare was the norm in the GDR. These differences resulted in substantially higher female labor force participation rates in the East than in the West, earlier returns to work after having children, and a higher number of hours worked for women, among other things.[7]

In this study, we use these institutional differences and the resulting differences in the gendered nature of work as the building blocks of an instrumental variable approach. Specifically, we rely on stark differences in occupational female shares between the GDR and West Germany across all occupations and the higher female labor force participation in the GDR. One might suspect that the stronger emphasis on gender equality in the political and institutional set-up of the GDR than in the West resulted in a more equal distribution of the two genders across occupations. However, that was not the case, as Rosenfeld and Trappe (2002) document. Occupations that in West Germany and other market-based economies were predominantly male-dominated, especially in manufacturing, were more integrated in the GDR, and women were employed in an overall wider range of occupations. At the same time, women in the GDR were more concentrated in occupations that were predominantly female-dominated in West Germany, especially in the service sector. In fact, the GDR had an overall slightly higher level of aggregate occupational segregation. Rosenfeld and Trappe (2002) cite different factors for the observed

[6] Constitution of the GDR, Article 24 (2), 1968.

[7] Rosenfeld et al., 2004, summarizes the empirical evidence on the differences in labor market outcomes in the GDR and West Germany. For example, women's labor force participation rate was 89% (men's was 92%) in the GDR 1989, whereas West German women's participation rate was 56% (men's was 83%). The gender pay gap at the time was equally large in the GDR and West Germany. Recent work that has investigated the consequences of the differences in the gendered nature of work of the two parts of Germany includes Beblo and Goerges, 2018, Lippmann et al., 2020, and Boelmann et al., 2024.

differences in occupational female shares, including gender ideology, differential responses to issues of labor scarcity (male guest workers in the West, women in the East), family policy, industrial structure, as well as educational and vocational training systems. Our instrumentation relies on the resulting differences in gendered occupational employment structures in the East and the West at the time of the collapse of the Berlin Wall in 1989, which we document below.

3 Data and Variables

3.1 Data Sources

Our main analysis relies on two data sets: the Qualification and Career Survey (QCS) and the German social security records, specifically the Sample of Integrated Labour Market Biographies (SIAB).

The SIAB of the Institute for Employment Research (IAB) is our primary sample. It is a 2% random sample of all employees subject to social insurance contributions in Germany; the data do not include civil servants and the self-employed. It provides high-quality information on gross daily earnings of workers, among other things. In the context of this study, one important advantage is that its large sample size leads to a positive number of observations even in fine age-occupation-gender cells. In addition, it is a panel data set that allows us to include individual and employer fixed effects to control for unobserved time-invariant person and employer characteristics.

The SIAB has however two well-known disadvantages: First, it contains gross daily earnings and a part-time dummy, but not hours worked. Thus, we cannot calculate hourly wages. As the part-time share of women in Germany is relatively large, and the hours distribution conditional on working part-time is wide, we focus on full-time working women and men in our main analyses.[8]

Secondly, in the SIAB, it is challenging to identify citizens of the GDR who moved from East to West Germany before 1992, the year in which East Germany was fully integrated into the social security system's data collection.[9] Our main wage regressions are for citizens of the FRG ("West Germans") only, so we need to separate them from East Germans who moved to the West after the Wall fell. To do the latter, we follow Boelmann et al. (2024) and others and, in a first step, focus on individuals residing in the West. We then exclude individuals we identify as East German based on the timing and place of their first appearance in the data.[10]

[8] Note that the results show similar patterns when we include part-time workers, as shown exemplarily for the basic specification in the Appendix Table C.2, Columns (3) and (4).

[9] The SIAB does not contain information on whether an individual grew up in the GDR or other information allowing the identification of former GDR citizens.

[10] More specifically, we classify those individuals as East German who entered the SIAB in 1992 in an

For the construction of the instrumental variables, we use the QCS. This is a representative cross-sectional survey of the German population by the IAB and the German Federal Institute for Vocational Training (BIBB). We use the survey waves from 1991/92 and 1998/99, i.e., the earliest sample periods after reunification, with about 20,000 observations each.[11] The survey asks respondents about their current job, education, training occupation, and, importantly, whether they grew up in the GDR or the FRG. In addition, for all individuals growing up in the East, regardless of where they reside after reunification, we know the occupation they were trained in at some point in the GDR. We use this information to construct our instrument, as explained in Section 4. The QCS is a survey of all labor force participants, i.e., it includes information on the employed and the unemployed. As a result, when constructing the instrument, we do not face the typical selection issues of many studies that only observe part of the relevant population (see discussion below).

We also use the QCS data when investigating the effect of changes in the occupational female shares on the tasks that workers perform at the workplace, as task information is not included in the SIAB data. For the analysis of the effects on occupational prestige, we rely on different data that we will discuss below.

3.2 Estimation Samples

Main Sample For our primary analyses based on SIAB data, we use the years 1985, 1986, 1991, 1992, 1998, and 1999, as these correspond to the years for which we have the QCS data that we use to construct the instrumental variables (see below). We use June 30 as the reference date to assign to an individual all relevant characteristics (e.g., occupation, age, employer) for the respective year.

Our analysis focuses on individuals with a vocational training degree. Vocational training is highly standardized in Germany and is part of the formal education system. The focus on this group of workers ensures that we compare men and women with very similar occupation-specific skills, with certificates documenting the type of occupational skills the degree holder acquired. In addition, we observe both the current occupation of this group of workers and their training occupation. The latter is crucial information in the construction of the instrument. This group also represents the largest group of German workers, encompassing around two-thirds of all employees.

The primary sample contains full-time working individuals aged 25 to 54. It covers

East German firm or 1991 in an unclassified firm (individuals were added in 1991, but the firm reporting structure in East Germany was only implemented in 1992), or who enter in West Germany between 1989 and 1991 and for whom we assume based on their age and education that they must have worked in East Germany before. More details on this issue as well as the general treatment of the SIAB data can be found in Appendix A.

[11] The surveys refer to two years, e.g., 1991/92, because they were launched in the winter months that span the two respective years.

774,803 male and 369,014 female West German employees in the West German labor market, reporting all the relevant information. Using administrative data, our sample is much larger than in related studies that often rely on survey data.[12]

Construction of Female Share One important conceptual decision relates to the appropriate level of aggregation for the female share. Taking the "devaluation" hypothesis from sociology literally, one would construct the female share on the occupation-time level. However, keeping with the economic literature, we use the age-occupation-time level in the main analyses. From Blau et al. (2013), we know, for example, that the primary mechanism that fosters occupational gender integration is the entry of new cohorts of women (i.e., younger women). In addition, similar to workers with the same level of education but different ages being imperfect substitutes in production (e.g., Card and Lemieux, 2001), the evidence suggests that workers in the same occupation but with different ages are imperfect substitutes in production (e.g., Prantl and Spitz-Oener, 2020). Finally, and as discussed in detail again in Section 4, constructing the female share on the age-occupation-time level allows us to include occupation-time fixed effects in the empirical analyses. These fixed effects capture occupation-specific unobserved factors that might change over time and that may be correlated with wages and the female share (such as changes in productivity owing to technological developments or amenities). So, already in the OLS case, we can account for one of the important endogeneity concerns one might have in the context of the research question.

Event-Study Sample For analyzing potential mechanisms of the estimated effect in Section 5, we also use an event-study design. For this design, we construct a continuous sample based on the SIAB data and ranging from 1984 to 1999. Only individuals who are observed for two consecutive years are included, i.e. if an individual is present in 1988 and 1989 but no other years, they will be included with their 1989 observation. The year 1984 is thus only used as a reference year and does not directly enter the sample. All other sample restrictions are implemented as above. The female share and migrant share instruments are now defined as the change in the female share in the respective age-occupation cell between 1985/86 and 1991/92, i.e., between waves 1 and 2 of the main sample. The dependent variable is the year-on-year change in log wages. Additionally, the sample is further divided between occcupation stayers - those who report in year y that they are in the same occupation as in year $y-1$ - and switchers, i.e., those who report being in occupation o in year y but in occupation o_{-} in year $y-1$.

3.3 Main Variables

Wages. The dependent variable in the empirical models is the logarithm of the real

[12] In compliance with the data confidentiality rules of the data provider, age-occupation-gender-time cells with fewer than 20 observations are excluded from the sample in all figures. Still, all observations are included in the regressions.

daily gross wage for native West German employees in the West German labor market, measured in Euros at prices in 2015.[13] In the main sample, men's mean real daily wage is 108.44 Euros in 1985/86, 118.45 Euros in 1991/92, and 116.01 Euros in 1998/99. The corresponding mean wage for women is 78.39 Euros in 1985/86, 88.61 Euros in 1991/92, and 90.25 Euros in 1998/99 (see Table 1). Over time, women's mean real daily wages increased from 72.3 percent of men's mean real daily wages in 1985/86 to 74.8 percent in 1991/92, and 77.8 percent in 1998/99. The difference between the genders is less pronounced at the median, with 75.9 percent in 1985/86, 77.3 percent in 1991/92, and 80.9 percent in 1998/99.

Female Shares. The explanatory variable of main interest is the share of women in each individual's age-occupation-time cell in West Germany. We rely on 6 age groups (25-29, 30-34, etc.) and 42 occupations to construct the age-occupation cells for each time period. Table 1 shows the summary statistics. In 1985/86, the mean female share for women was 64 percent, compared to 23.2 percent for men. At the end of our observation period, it was 63.5 percent for women and 26.1 percent for men. Thus, over time, the occupational female share among women decreased by about 0.5 percentage points, while that of men increased by about 2.9 percentage points. In addition to evolving in opposite directions, the absolute change in the female share for average male employees was larger than for female employees.

Migrant Shares. To control for the general supply shock that hit the West German labor market from East-West migration after the fall of the Wall in 1989, we construct the *migrant share* variable. This captures the share of workers of East German origin in West Germany in each age-occupation-time cell. As shown in Table 1, there were no East German migrants in West Germany in 1985/86. The mean migrant share increased quickly to 7.42 percent for men and 6.05 percent for women in 1991/92.

Instrumental variables. We discuss the construction of the instrumental variables in the next section.

4 Empirical Analyses

4.1 Empirical Specification and OLS Results

We estimate the following individual-level wage regressions:

$$\begin{aligned} \log w_{iaot} = \beta_0 + \beta_1 f_{aot} + \beta_2 m_{aot} + \beta_3 \text{female}_i \\ + \varrho_{ao} + \varsigma_{ot} + \tau_{at} + \theta_{ao\times\text{female}} + \eta_{at\times\text{female}} + \nu_{iaot} \end{aligned} \tag{1}$$

The subscripts indicate individuals (i), age groups (a), occupations (o), and years

[13] We impute top-coded wages following the methodology proposed by Dauth and Eppelsheimer (2020).

(t). The variable of primary interest is f_{aot}, the fraction of women in each individual's age-occupation-time cell, defined as $f_{aot} := L^F_{aot}/(L^F_{aot} + L^M_{aot})$, where L^F_{aot} and L^M_{aot} denote the number of women and men, respectively, in a given age-occupation-time cell.

The indicator variable female$_i$ identifies female individuals. We control for age-occupation fixed effects (ϱ_{ao}), so the coefficient on the female share variable, β_1, is identified from within age-occupation cells over time when estimating with OLS. The specification also controls for occupation-time fixed effects (ς_{ot}) to account for changing occupation-specific characteristics (e.g., technological change, demand shifts); and age-time fixed effects (τ_{at}) to capture life-cycle wage patterns that vary over time. We further include gender-specific fixed effects by interacting the female dummy with both age-occupation ($\theta_{ao} \times$ female$_i$) and age-time ($\eta_{at} \times$ female$_i$) fixed effects. This allows for differential wage patterns between men and women across age, occupation, and time.

To account for labor supply shocks due to German reunification, we include m_{aot}, the share of East German migrants in the age-occupation-time cell. Since both f_{aot} and m_{aot} are potentially endogenous, we later instrument for them, as described below.

The basic set-up of Equation 1 closely mimics specifications often estimated in immigration studies following the skill-cell approach as pioneered by Borjas (2003) and variants thereof that use occupations instead of education as skill category (e.g., Prantl and Spitz-Oener, 2020). It fully captures the wage effects of the baseline labor supply shock (m_{aot}). It considers various unobserved factors through the various fixed effects, as explained in the previous paragraph. The main difference in our specification is the inclusion of the variable that captures the impact of occupational female shares (f_{aot}) on wages. Therefore, the effect we find for β_1 is net of any potential effects typically considered in the relevant migration literature, including the baseline labor supply shock to the cell.

Table 2, column (1), shows the results for β_1 and β_2 when estimating specification (1) using Ordinary Least Squares (OLS) and our primary sample for estimation.[14] The results suggest that changes over time in the share of women in an age-occupation cell negatively correlate with wages in the same cell. In addition, the coefficient on the migrant share variable is also negative and statistically significant, indicating that the labor supply shock in the context of German reunification put downward pressure on West Germans' wages.

[14]Note that in all results, we show standard errors clustered at the age-occupation-time cell level to allow for unrestricted correlations between the observations within these cells.

4.2 Instrument Construction, Reduced Form, and First Stage Results

With the set of controls in Specification (1), we already control for unobserved changes in occupational characteristics that might correlate with female shares and wages (e.g., amenities). However, any endogeneity concerns one might have related to unobserved time-varying factors correlated with age-occupation female shares and wages still threaten the causal interpretation of our estimates. For example, one might be concerned about employers aiming to attract young women into male-dominated occupations in which labor scarcity is particularly imminent by offering amenities that are especially attractive to young women. In the next step, we therefore apply an instrumental variable (IV) strategy.

The IV approach is based on the increase in the potential pool of German workers owing to the unexpected Fall of the Berlin Wall on November 9, 1989, the associated breakdown of the GDR, and the subsequent merging of the labor markets of the "two Germanies". The term 'potential pool' indicates that we consider East Germans who lived in West or East Germany during the 1990s when constructing the instruments (i.e., East-West migrants and East Germans who stayed in the East): all East Germans could potentially work in West Germany after reunification, whether they chose to do so or not, but all of them were prevented to work in the West prior to reunification. The instrumental variables use the pre-1989 occupational fields of training of East German men and women residing in East or West Germany in 1991/92.[15]

The instrument is constructed in two steps. In Step 1, we construct the gender-specific pool of East Germans per age-occupation-time cell $L^*_{g,aot}$:

$$L^*_{g,aot} = \begin{cases} \sum_{k=1}^{K} \omega_{aokt} L_{gakt} \quad \text{if} \quad t = 1991/92 \\ \\ \text{artificial aging of 1991/92 values if t=1998/99} \end{cases} \tag{2}$$

$$\omega_{g,aokt} = \begin{cases} 1 \quad \text{if} \quad k = o \\ \\ \frac{L_{gaokt}}{L_{gaot}} \quad \text{if} \quad k \neq o \end{cases} \tag{3}$$

The subscript g indicates the gender ($g :=$ female (F) or male (M)), and k indicates the occupation in which the individual was trained during GDR times, i.e., before 1989. As Equation 2 illustrates, the pool-measure is a weighted sum of the number of East German men or women in 1991/92 over all training occupations. The weights ω_{aokt} are

[15] In the main specification, we rely on this earliest possible sample period after reunification to exclude endogenous labor market participation decisions to the best extent possible. The results are robust to also using data from 1998/99 rather than implementing artificial aging, as we do here.

1 if the pre-1989 training occupation k equals the current occupation o (see Equation 3). They are equal to the share of West Germans in 1985 with training occupation k in occupation o if the training occupation is unequal to the current occupations.[16] The pool measure thus indicates the number of East Germans by gender who are potentially available to work in a specific age-occupation cell.

For the measures in 1998/99, we "age" the variables generated for 1991/92 artificially to 1998/99. Thus, variation in the instrument during the 1990s is not driven by any change in the labor force participation of East Germans over time, which is potentially endogenous. Any variation in $L^*_{g,aot}$ between 1991/92 and 1998/99 arises from ageing of the East German workforce of 1991/92.

In Step 2, we take the pool measures to construct the instrumental variable for f_{aot}, the fraction of women in each individual's age-occupation-time cell in West Germany, as well as the instrumental variable for the migration shock measure m_{aot}. Specifically, the instrument for f_{aot} is defined as follows:

$$f^{IV}_{aot} = \begin{cases} \left(\frac{L^*_{F,t}+L^{West}_{F,pre}}{L^*_{F,t}+L^*_{M,t}+L^{West}_{pre}}\right)_{ao} & \text{if} \quad t = 1991/92, 1998/99 \\ \left(\frac{L^{West}_{F,pre}}{L^{West}_{pre}}\right)_{ao} & \text{if} \quad t = 1985/86 \end{cases} \tag{4}$$

with pre referring to the years 1985/86.

As Equation 4 illustrates, the instrumental variable varies on the age-occupation-time level. For the years before 1989, it is constructed as the female share in West Germany in 1985/86 per age-occupation cell ($L^{West}_{F,pre}$ indicating the number of women, and L^{West}_{pre} indicating the total number of employees, 'pre' referring to the fact that this measure is calculated before 1989). Thus, for the years before 1989, the instrument is equal to the actual female share since the pool of East German workers who could potentially work in the West is 0 before reunification. We add the potential pool of East German workers to the 'pre' values for the years after reunification. In the numerator, we add the potential pool of relevant female East Germans; in the denominator, we add the potential pool of relevant male and female East Germans. By adding the pool of potential East German workers to the actual number of West German workers prior to reunification, we implicitly also take into account the different sizes of the two groups.[17]

[16] E.g., for the occupation of carpenters (o=carpenter), the labor supply pool would count all (male and female) East Germans who were trained as carpenters ($k = o$), and add to that, for example, 20% of the East Germans trained as plumbers (k=plumber), in case that 20% of all carpenters among West Germans were trained as plumbers. The calculation of the weights thus relies on data from West Germans and uses data from 1985, i.e., many years before the Wall fell in 1989.

[17] For example, even if the female shares in an occupation-age cell in East and West were very different prior to reunification, German reunification could barely alter the West German female shares if the number of East Germans in this specific occupation-age cell was very small.

As discussed in Section 2, differences in gender ideology and family policy, different reactions to issues of labor scarcity, differences in industry structure, as well as differences in educational and vocational training systems in the "two parts of Germany" produced stark differences in the gendered nature of work in the GDR and FRG before 1989. Most prominent is the much larger labor market attachment of East German women than West German women. Importantly for this study, the distributions of men and women across occupations also differed in the East and the West.

Figure 2 illustrates the variation in the gender composition of occupations that we rely on in our instrumental variable approach (to reduce the complexity of the figure, we show the variation on the occupation level here). On the x-axis, occupations are ranked based on the occupation size in the West in 1985/86 (measured in number of employees) from large to small. The y-axis shows the percentage point differences in the female share between the East and West before reunification.[18]

Consistent with East German women's overall higher labor force participation, the East-West difference in the female share is generally positive. The largest occupation is office workers, where the female share in the East is almost 30 percentage points larger than in the West. For bookkeepers, the difference is the largest, at over 70 percentage points. There are only two occupations in which the female share was larger in the West: household and building cleaners, and metal polishers and connectors.[19]

Figure 3 graphically illustrates the first stage. The x-axis shows the change over time (from 1985/85 to 1991/92 and from 1991/92 to 1998/99) in the female share instrument per age-occupation cell, and the y-axis shows the change in the corresponding actual female share in West Germany. The observations are grouped into 60 bins, each including about 12,900 individuals, with the red line representing a linear fit. The slope coefficient is 0.215, and the relationship is highly statistically significant and not driven by outliers. The instrument appears to be a valid predictor of the endogenous variable.

Equation 5 shows the formula that we use to construct the migrant share instrument (m_{aot}^{IV}) with which we account for the fact that German reunification was also a labor supply shock, not only a shock to female shares:

$$m_{aot}^{IV} = \begin{cases} \left(\frac{L^*_{F,t}+L^*_{M,t}}{L^*_{F,t}+L^*_{M,t}+L^{West}_{pre}}\right)_{ao} & \text{if} \quad t = 1991/92, 1998/99 \\ 0 & \text{if} \quad t = 1985/86 \end{cases} \tag{5}$$

[18]For the West, we use the actual female share in an occupation in 1985, i.e., $f_{pre,ao}$ aggregated to the occupation level. For the East, we calculate $L^{East}_{F,pre}/L^{East}_{pre}$ on the level of training occupations during GDR times.

[19]We show detailed summary statistics on the instruments in Appendix Table C.1. The values of the female share instrument range from 2.77 percent at the 25^{th} percentile for men in 1985/86 to 78.08 percent for the 75^{th} percentile of women in 1991/92, closely mirroring the dispersion present in the actual female shares.

with pre=1985/86.

The main components of the instrument are again the pool measures defined in Step 1 above, and the reasons for its exogeneity are the same as for the female share instrument. Figure 4 graphically shows the first stage in the same setup as for the female share. The slope coefficient of the fitted line is 0.185.[20]

Table 3 shows the results of the first-stage regressions for the female share (Column 1) and the migrant share (Column 2). The corresponding reduced form results are shown in Column (3). All instruments are positive and highly significant in the relevant specification, i.e., the coefficients on the diagonal of Columns 1 to 2 drive the variation in the respective instrumented variables. In addition to the standard F-test statistic, we report Sanderson-Windmeijer (SW) F-test statistics that account for the multiple instruments setting (Sanderson and Windmeijer, 2016). The values do not indicate that the instruments correlate weakly with the endogenous regressors. The reduced form regression shows a statistically significant, negative association between wages and the female share instrument.

4.3 Second Stage Results

Column (2) of Table 2 shows the second-stage IV estimates. First, note that the baseline labor supply shock brought about by German reunification led to lower wages, as the coefficient on the migrant share variable (m_{aot}) at the bottom of the table is negative and highly statistically significant.[21] The IV estimate for our primary variable, the female share shock (f_{aot}), in the first row of Column (2) indicates that an increase in the female share causally leads to lower wages. The previous OLS results seem not to be driven by, for example, employers aiming to attract young women into male-dominated occupations in which labor scarcity is imminent by offering amenities that are especially attractive to them. Because we use instrument variation for identification, include the baseline labor supply shock in the context of German reunification (m_{aot}), and also add the various fixed-effects, the results suggest that it is the gender composition itself that has an independent effect on wages.

Comparing the OLS and IV estimates suggests that instrumenting is important.

[20]As a robustness check, in Figure B.1, we present similar first stage figures for the female and migrant share, in which we additionally control for the respective other variable. This specification is closer to the actual first stage used in the regression. The slope coefficients remain almost unchanged, showing that our instruments are valid predictors for the respective endogenous variables. To further illustrate the dynamics of the instrument, we also show the first stage graphs separately for 1985/86-1991/92 (left panel) and 1991/92-98/99 (right panel) in Figure B.2 of the Appendix. For the female share, it becomes clear that most of the variation comes from the first period. Analogously, in Columns (1) and (2) of Table C.2, we repeat the main analysis using only the years 1985/86 and 1991/92.

[21]This result is very similar to the result reported in Prantl and Spitz-Oener (2020). Following Borjas (2003), the wage elasticity, evaluated at the increase in the ratio of East to West Germans between 1985/86 and 1998/99 (6.93), is $-1.190/(1 + 0.0693)^2 = -1.04$.

Specifically, the change in the estimated coefficients suggests that female shares increased in thriving age-occupation cells, in which demand for labor and, thus, wages were growing. Of course, measurement error and the resulting attenuation bias towards zero in the OLS regression are also consistent with the increase in the absolute size of the coefficients between the OLS and IV specifications.

What is the quantitative importance of the negative effect of the female share on wages? The 2SLS results in Column (2) imply that a one percentage point increase in the female share in the age-occupation cell is associated with 0.7 percent lower wages.[22] Evaluated at the mean change in female shares between 1985/86 and 1998/99, which was 3.63 percentage points, this implies 2.51 percent lower wages due to the increasing female share.

5 Mechanisms

We use a series of heterogeneity analyses to pinpoint the mechanisms driving our key results.

Evidence on local level. First, we consider whether the effect of changes in the female share on wages differs by localized developments in occupational female shares. Specifically, we consider district heterogeneity (relying on 325 West German *Kreise*) in Column (3) by allowing the effect to differ by districts in which the local female share in the respective occupation actually increased (the interaction effect in row 2) and districts in which the female share in the respective occupation did not increase (the baseline effect in row 1). The results show strong negative wage effects also in districts where the female share didn't increase; the effect is only marginally (5%) more negative in districts where the female share actually increased. These results indicate that wages are affected by the perceived feminization of an occupation rather than the actual feminization on the local level. They corroborate the notion that the negative wage effects in feminizing occupations are not driven by changes in the actual content of work, but by who is perceived to be doing the work.

Direct evidence on prestige. To provide more direct evidence on the changes in the prestige of occupations and their task content, we rely on survey data. Occupational prestige scores are available based on surveys conducted in 1979/1980 in West Germany and in 2017/2018 in unified Germany. The details of the data are discussed in Appendix A.5. While the survey from 1979/1980 is very well suited to capture the pre-unification prestige of occupations, we would have preferred to have survey information in the post-unification period that better matches the relevant period of our study. Unfortunately,

[22] Note that, in the regression, the female share is defined as running from 0 to 1, but in the interpretation, we use the more intuitive formulation of running from 0 to 100.

no such survey has been conducted.

To provide suggestive evidence that prestige decreased in occupations that experienced increases in the female shares, we compare the evolution of prestige scores in occupations with a high versus a low female share increase, similar to the analysis in Figure 1. In Figure 5, we show the average prestige score in the early period (1979/1980) and in the late period (2017/2018), separately for the two groups of occupations that we also consider in Figure 1, i.e., "Top 40%" and the "Bottom 40%". We normalized the scores to one for the early period. The group with the small female share increase (Bottom 40%) experiences a strong increase in prestige over time, whereas the score drops in the group with the high female share increase (Top 40%). Taken together, the prestige of the group of occupations with a large female share increase deteriorated sharply relative to the group with a small increase, suggesting that increases in the occupational female share are associated with a devaluation of the prestige of occupations.

Additionally, in Figure 6, we show binned scatter plots of the changes in the female share and in occupational prestige. We calculate the change in the prestige score between the early and the late period at the occupation level and plot this against the change in the female share (graph on the left) or the female share instrument (graph on the right) between 1985/86 and 1991/92. The change in the prestige score shows a clear negative correlation with both the change in the female share and the change in the female share instrument, with correlation coefficients of -0.24 and -0.13, respectively. Overall, the results are highly consistent with the "devaluation" hypothesis in sociology.

Direct evidence on task changes. The QCS data has the advantage of including information on the tasks that workers perform at the workplace. This information has been used intensively, for example, in Spitz-Oener (2006) and Black and Spitz-Oener (2010). Table 4 shows the IV results of regressions that estimate Equation 1 using the QCS data. The outcome variables are task shares on the worker level, indicating the share of analytical (interactive, routine cognitive, routine manual, and non-routine manual) tasks among all tasks a worker performs at the workplace.[23] By and large, the results suggest that changes in the female share (and more generally, changes in the migrant share) have no impact on the tasks workers perform at the workplace. The coefficients on the female share variable are negative and weakly statistically significant (at the 10 % level) for the analytical and interactive task shares. However, the quantitative importance of the estimated effects is very small and unlikely to drive the effects on wages we find.[24]

[23]The corresponding OLS regressions are shown in Table C.4, and the corresponding first stage regressions are shown in Table C.5. More details on the construction of the sample can be found in Appendix A.6.

[24]The average increase in the female share between 1985/86 and 1998/99 was 3.63 percentage points. The mean task shares during the 1990 are 9.1% for the analytical tasks category, 32.6 % for the interactive task category, 32.2 % for the routine cognitive, 25.7 % for the non-routine cognitive, and 35.7% for the non-routine manual task category. Evaluated at the mean, the coefficient on the female share in the first

Additional mechanisms. In Column (4) of Table 2, we extend the baseline specification by allowing the effect of occupational feminization to vary for across worker subgroups. Specifically, we interact with the female share variable, f_{aot}, indicators for inflows (new entrants to an occupation) and outflows (those exiting an occupation).[25] First, note that the coefficient on the main variable remains large and highly statistically significant. However, the coefficient on f_{aot} × Inflow is negative and highly significant (coef.: –0.14, Std dev.: 0.014), indicating that the wage decline associated with increasing feminization is particularly pronounced for individuals entering an age-occupation cell. This could be a selection effect, with lower skilled individuals entering the occupation, or could indicate a higher wage flexibility among individuals new to an occupation. In contrast, the small and insignificant coefficient on f_{aot} × Outflow suggests that feminization affects those leaving an occupation similarly to those who stay. If the wage penalty were due to task downgrading or a decline in occupational skill requirements as more women enter (e.g., "feminized jobs become easier or less demanding"), then we would expect the effect to be concentrated among those who remain in the occupation, not those who leave. However, since both groups are similarly affected, this points to a mechanism based on external perceptions, and, again, is consistent with the "devaluation" hypothesis in sociology.

The latter result is corroborated by the results from an event study. This design allows us to investigate the dynamic effects of the female share shock on age-occupation cells in the context of German reunification. In addition, we examine the sensitivity of our results to workers' occupational mobility.

Specifically, we estimate the following specification:

$$\Delta \log w_{iao} = \gamma_1 \Delta \hat{f}_{ao}^{91/92-85/86} + \gamma_2 \Delta \hat{m}_{oa}^{91/92-85/86} + a_i + \epsilon_{iao} \qquad (6)$$

As before, the subscripts indicate individuals (i), age groups (a), and occupations (o). $\Delta \log w_{iao}$, the change in log wages, is calculated for each i for consecutive years (i.e., 1984-85, 1985-86, etc., up to 1998-99). To construct this outcome variable, we adjust the estimation sample so that we only keep observations observed in both years for each pair of two consecutive years. Therefore, we hold the composition of workers constant for two

column suggests a decline in the analytical tasks shares of 0.43 percentage points, which accounts for 4.7% of the mean task share. The corresponding figure for the interactive task share is 0.01%.

[25] We define inflows into and outflows from occupations with reference to the preceding or following wave, respectively. More concretely, individuals are defined as an inflow to an occupation o if they are in occupation o in the current observation wave but were not in this occupation in wave $t-1$. The individual can thus flow into occupation o from a different occupation o_{-} or from outside the sample. Outflows are constructed in a symmetrical manner, i.e., individuals are considered outflows from occupation o if in the following wave $t+1$, they report a different occupation o_{-} or do not appear in the sample altogether. To obtain well-defined flow variables also for the first and last waves of the sample, respectively, we add data from the years 1980 and 2004 as reference points. The data are prepared using the same sample restrictions as for the main estimation sample. These observations are only used to define in- and outflows, but are discarded again before the estimation itself.

years. The female share shock is constructed as follows: $\Delta f_{ao}^{91/92-85/86} = f_{ao}^{91/92} - f_{ao}^{85/86}$. The relevant instrumental variable is constructed as follows: $\Delta f_{oa}^{IV,91/92-85/86} = f_{ao,91/92}^{IV} - f_{ao,85/86}^{IV}$. We estimate annual regressions of wage changes between the consecutive years from 1984 to 1999, keeping the female share shock variables consistently defined from 1985/86 to 1991/92. We then investigate the dynamic effects by summing over the respective coefficient estimates. In the latter, we use 1989, when the Berlin Wall fell, as the starting point and add up the coefficients backward and forward.

Figure 7 shows the results for two samples. First, "all" (the black line) includes all West German workers aged 25-54 between 1984 and 1999, who are observed in both years for each pair of two consecutive years. Secondly, "stayers" (the blue line) includes only the subsample of those who remained in the same occupation during each pair of two consecutive years. The figure shows that, before 1989, the estimated coefficients for both samples are small and not statistically different from zero (the grey shadow shows the 95% confidence interval for the black line and the blue shadow for the blue line). There are no pre-trends, so the results corroborate that the age-occupation cells hit by the female share shock after 1989 were not affected by other forces before 1989 that might hamper causal effect identification. After 1989, wages decline quite sharply for two years for both "stayers" and "all". They continue to fall for "all" but flatten out for those who stayed in the same occupation. As indicated by the two confidence intervals, the results for the two samples are not statistically different.

The similar patterns for both "stayers" (those who remain in the occupation) and "all" (which includes movers) show that the work content does not need to change for wages to fall if female shares increase. If task changes or declining skill demands drove the wage decline, we'd expect it to be stronger for stayers. But since it isn't, the perception of the occupation itself appears to change, not the content. These results also indicate that the main results are not exclusively driven by high-skilled individuals leaving feminized occupations, or low-skilled individuals entering feminized occupations, i.e. they are not driven by selection.

Column (5) of Table 2 includes an interaction between the female share variable and an indicator for workers older than 39. It could be that the workers who enter an occupation are mostly young and less experienced. The more negative effect for the entrants could reflect their lower skill levels rather than the entry effect. The change in the specification has a minimal impact on the coefficient for entrants. The coefficient on the interaction term for older workers is large, positive and statistically significant, suggesting that older workers experience a less negative wage response, or no wage response at all, to occupational feminization. This may reflect greater wage rigidity or seniority protections among older workers, shielding them from downward wage pressure.

In Column (6), we show the results of a specification in which we test for non-linearity in the wage effects, depending on the initial female share in the occupation-age cell. In this specification, we interact the female share variable with dummies, indicating occupation-age cells with female shares between 25 and 50 percent, 50 to 75 percent, and larger than 75 percent in 1985 (the baseline being occupation-age cells with a female share below 25 percent, i.e., the most male-dominated cells). The results suggest that the wage effects are uniform across the distribution. The effect on wages is negative for the initially most male-dominated occupation-age cells (the baseline), and the coefficients on the interaction terms are positive, but not statistically significant and neither statistically significantly different from each other.

Heterogeneity by establishment characteristics and economic environment. In Table 5, we extend our baseline specification with interaction terms to separately identify the effects of changes in the occupational female share on wages for different types of establishments and differences in establishments' economic environments. Column (1) shows the difference by establishment size, with the negative effect of the occupational female share on wages being almost 50% larger for workers in small establishments.[26] The differential effect by establishment size is consistent with the role of works councils in the German labor market. Works councils are key to the German system of establishment-level co-determination. They are supposed to foster labor-management cooperation by providing incumbent workers with a strong influence on firms' decision-making. Large establishments are more or less universally covered by works councils, whereas this is not the case for small firms.[27] The larger adverse effect for workers in small than large establishments is consistent with works councils mitigating the negative wage impacts on workers of increases in female shares. It is also consistent with the devaluation hypothesis, as the wage setting is less formalized in firms without work councils. These institutional differences can make it easier for changes in societal perceptions of occupational value to affect wages in small firms than in large firms. Or, put differently, institutions that formalize the wage-setting process can moderate how occupational devaluation manifests itself in wages.

Next, we investigate the potential heterogeneity of the effects depending on the economic environment, starting with the competitiveness of the product market in which the firms operate. We know that discrimination is less prevalent in competitive environments.[28] In Column (2), we show the results of a specification in which we test whether product market competitiveness affects the effects of changes in the occupational female

[26] We define establishments with less than 50 employees as small. Although not reported in the table, the specification includes the main effect for small establishments, which captures any wage difference associated with establishment size.

[27] Addison et al. (1997), for example, use data from 1994 to show that about 93% of the employees in West German establishments with fifty or more employees were represented by works councils, but there were only 18% in establishments with fewer than fifty employees.

[28] The leeway to discriminate is considerably restricted in competitive markets, see Becker (1957). The seminal paper by Black and Strahan (2001) provides empirical evidence by showing how deregulation in the US banking sector reduced discrimination against women.

share on wages.[29] The negative effect is more than 25% larger in an environment that restricts competition compared to the competitive environment. The finding suggests that wages may reflect more freely societal biases and, therefore, allow for a devaluation in environments with restricted competition, in which market forces are weak and wages deviate from the marginal product of workers.

In Table 5, Column (3), we show results of a specification in which we test for differential effects in the service and manufacturing sector. The results suggest that there are no such differences. In Column (4), we refine the specification by distinguishing between low and high-skilled services (a detailed definition of all these variables can be found in Appendix A). The results indicate differential effects within the services sector, with low-skilled services being somewhat more negatively affected by the increase in the female share than the manufacturing and high-skilled services sectors. However, the quantitative importance of this difference is minimal.

In Column (5), we take a closer look at the employers and their pay schemes, i.e. whether they are high- or low-paying employers. Firm-specific pay schemes are important in men's and women's wages and gender wage gaps.[30] We adjust the specification to investigate whether wages react differently to changes in occupational female shares in firms with different pay schemes. In particular, we differentiate three types of employers: high-, medium-, and low AKM fixed effects employers.[31] We do find statistically different effects depending on firms' pay schemes. For the high-paying firms, i.e., employers in the highest tercile of the AKM fixed effects, the negative wage effects of increases in the female shares are the smallest; they are the highest for the low-paying firms. The difference between the high and medium-AKM fixed effects firms is quantitatively small, but with 33%, the effect is sizable different from the baseline for low-AKM fixed effects firms. The finding is consistent with the devaluation hypothesis and institutions playing a mitigating role in the translation of female shares into observed wages. Low AKM fixed-effects firms are typically small firms with less formalized wage-setting processes than larger firms with high AKM fixed-effects.

In Column (6), we allow for differential effects of the increase in the female share depending on employers' degree of monopsony power in an occupation. The recent literature has abundantly documented that firms have monopsony power in the labor market and often can pay workers below their marginal revenue product.[32] We measure labor market

[29]Following Prantl and Spitz-Oener (2020), we rely on the German Trade and Crafts Code that restricts the entry of firms to distinguish competitive and non-competitive product markets. Note that the specification also includes the dummy that indicates the competitive economic environment with no entry restrictions separately, thereby accounting for all unobserved factors that might influence wages differentially in the different economic environments.

[30]See, e.g., Card et al. (2016) and, for Germany, Bruns (2019).

[31]Based on the AKM firm fixed effects provided by the IAB, we classify firms into terciles based on those effects. High AKM firm fixed effect employers are those with fixed effects in the highest tercile, and so on. A detailed definition can be found in Appendix A.

[32]The chapter on monopsony power in the labor market in the forthcoming Handbook of Labor

concentration on the occupation-region-time level, computed as flow-adjusted Herfindahl-Hirschman indices (HHI) following Arnold (2021).[33] We group the occupation-region cells into low-, middle-, and high-concentration terciles for the heterogeneity analyses. Specifically, we interact the female share variable with dummies indicating medium and high concentration occupation-region measures.[34] The results suggest that the effects of increases in the female shares are relatively uniform. The coefficients on the interaction terms are statistically different from the main effect in Row (1) of Column (6); however, the quantitative importance of the effects is again small.

5.1 Gendered Effects

In Table 6, we show results of specifications that investigate whether the effect of increases in the female share in age-occupation cells on wages differs by gender. Both the OLS results in Column (1) and the IV results in Column (2) indicate that the effect for women is smaller than it is for men.[35]

To gain a deeper understanding of the sources of these differences by gender, we augment the specification with individual fixed effects in Column (3), which control for time-constant unmeasured individual productivity or taste differences that might be correlated with the gender composition of an age-occupation cell and wages, among other things. For women, relying on the within-individual variation for identification has a relatively large consequence on the result, with the composite coefficient not being statistically different from zero (coef.: -0.682 + 0.610 = -0.072; std. 0.102). The effect for men is also somewhat more muted than the result in Column (2); the size of the coefficient drops by about 15 percent, but it remains highly statistically significant. Overall, these results corroborate that men bear the adverse effects of increases in occupational female shares to a much larger extent. In contrast, the effect for women seems to be mainly driven by sorting effects (i.e., low-wage women sorting into high-female share cells). This result is corroborated in Column (4), where the specification includes firm fixed effects instead of individual fixed effects. The change in the coefficient when compared to the result in Column (2) again suggests that sorting is an important mechanism for explaining the wage effects for women, while it is not for men.

Economics provides an excellent review, see Azar and Marinescu (2025).

[33]For details, see Appendix A. For this measure, we rely on the information on commuting zones available for the social security records, among other things.

[34]As for the previous specifications, the regression also includes the main effects, but the results are not shown.

[35]The lower quantitative importance of the female share for women's wages is in line with findings in part of the literature (see, e.g., Harris, 2022).

6 Conclusion

In recent decades, progress in reducing the gender pay gap has slowed down or even stopped in many industrialized countries. Based on the observation that women continue to be overrepresented in low-paying occupations and men in high-paying occupations, many believe that reducing the degree of occupational segregation is necessary to reduce the gender wage gap further. This widely held view is also reflected in recent policies that aim to increase female labor force participation in highly-paying male-dominated fields, such as STEM. Proponents of those policies highlight that women would catch up to men's higher wages by moving into those highly-paying occupations, thereby reducing the overall gender pay gap.

In this paper, we provide causal evidence that an increase in the female share in an occupation leads to lower wages for men and, to a lesser extent, for women. We establish causality by exploiting a gendered labor supply shock caused by German reunification.

Overall, the results suggest that decreasing occupational segregation might be effective in closing the gender wage gap. However, the mechanism we identify is quite different than the one that is typically discussed. We show that in the process of reducing occupational segregation, wages change. In occupations in which the female share increases, wages decline in particular for men, which contributes to a reduction in the gender wage gap if occupational sorting decreases. We show that these adverse effects of increasing female shares are pervasive and all-encompassing, and not driven by changes in skill requirements or the task content in occupations. Our combined results are consistent with the "devaluation hypothesis" put forth in the sociology literature. They also help explain the reluctance of men to have women enter their occupation, as put forward in the "pollution" theory.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67 (2)*, 251–333.

Addison, John T., Claus Schnabel, and Joachim Wagner (1997). On the Determinants of Mandatory Works Councils in Germany. *Industrial Relations: A Journal of Economy and Society 36*(4), 419–445.

Arnold, David (2021). Mergers and Acquisitions, Local Labor Market Concentration, and Worker Outcomes. *R&R at American Economic Review*.

Azar, José, and Ioana Marinescu (2025). Monopsony Power in the Labor Market. In C. Dustmann and T. Lemieux (Eds.), *Handbook of Labor Economics, forthcoming.* Elsevier, Amsterdam.

Beblo, Miriam, and Luise Goerges (2018). On the nature of nurture. The malleability of gender differences in work preferences. *Journal of Economic Behavior & Organization 151*, 19–41.

Becker, Gary S (1957). Human Capital, New York 1964. *BeckerHuman Capital1964*.

Bertrand, Marianne (2020, May). Gender in the Twenty-First Century. *AEA Papers and Proceedings 110*, 1–24.

Black, Sandra E., and Alexandra 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.

Black, Sandra E, and Philip E Strahan (2001). The division of spoils: rent-sharing and discrimination in a regulated industry. *American Economic Review 91*(4), 814–831.

Blau, Francine D., Peter Brummund, and Albert Yung-Hsu Liu (2013). Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System. *Demography 50*(2), 471–494.

Blau, Francine D., and Lawrence M. Kahn (2017, September). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature 55*(3), 789–865.

Boelmann, Barbara, Anna Raute, and Uta Schoenberg (2024). Wind of Change? Cultural Determinants of Maternal Labor Supply. *American Economic Journal: Applied Economics Forthcoming.*

Böhm, Michael J., Hans-Martin von Gaudecker, and Felix Schran (2024). Occupation Growth, Skill Prices, and Wage Inequality. *Journal of Labor Economics 42*(1), 201–243.

Borjas, George J. (2003). The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics 118 (4)*, 1335–1374.

Bruns, Benjamin (2019). Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap. *American Economic Journal: Applied Economics 11*(2), 74–113.

Bárány, Zsófia L., and Christian Siegel (2018). Job Polarization and Structural Change. *American Economic Journal: Macroeconomics 10*(1), 57—-89.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics 131 (2)*, 633–686.

Card, David, Jörg Heining, and Patrick Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Card, David, and Thomas Lemieux (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics, 116 (2)*, 705–746.

Dauth, Wolfgang, and Johann Eppelsheimer (2020). Preparing the sample of integrated labour market biographies (SIAB) for scientific analysis: a guide. *Journal for Labour Market Research 54*(1), 1–14.

Ebner, Christian, and Daniela Rohrbach-Schmidt (2021). Das gesellschaftliche Ansehen von Berufen – Konstruktion einer neuen beruflichen Ansehensskala und empirische Befunde für Deutschland. *Zeitschrift für Soziologie 50*(6), 349–372.

England, Paula, George Farkas, Barbara Kilbourne, and Thomas Dou (1988). Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects. *American Sociological Review 53*(4), 544–558.

Felfe, Christina (2012). The Willingness to Pay for Job Amenities: Evidence from Mothers' Return to Work. *Industrial and Labor Relations Review 65*(2), 427–454.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter (2006). Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Journal of Applied Social Science Studies 126*, 405–436.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry (2022). Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D030.V10.0.

Frietsch, Rainer, and Heike Wirth (2001). Die Übertragung der Magnitude-Prestigeskala von Wegener auf die Klassifizierung der Berufe. *ZUMA Nachrichten 25*(48).

Goldin, Claudia (2014). A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings. In *Human Capital in History: The American Record.*, pp. 313–348. University of Chicago Press.

Greenberg, Kyle, Melanie Wasserman, and Anna Weber (2024). The Effects of Gender Integration on Men: Evidence from the U.S. Military. NBER Working Paper No. 33235.

Harris, Jorgen (2022). Do wages fall when women enter an occupation? *Labour Economics 74*, 102102.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica 87*(5), 1439–1474.

Levanon, Asaf, Paula England, and Paul Allison (2009). Occupational Feminization and Pay: Assessing Causal Dynamics USing 1950-2000 U.S. Census Data. *Social Forces 88*(2), 865–891.

Lippmann, Quentin, Alexandre Georgieff, and Claudia Senik (2020). Undoing Gender with Institutions: Lessons from the German Division and Reunification. *The Economic Journal 130*(629), 1445–1470.

Lochner, Benjamin, Stefan Seth, and Stefanie Wolter (2023). AKM Effects for German Labour Market Data 1985-2021. FDZ-Methodenreport, 01/2023 (en), Nürnberg.

Macpherson, David A., and Barry T. Hirsch (1995). Wages and Gender Competition: Why do Women's Jobs Pay Less? *Journal of Labor Economics 13*(3), 426–471.

Murphy, Emily, and Daniel Oesch (2016). The Feminization of Occupations and Change in Wages: A Panel Analysis of Britain, Germany, and Switzerland. *Social Forces 94*(3), 1221–1255.

Prantl, Susanne, and Alexandra Spitz-Oener (2020). The Impact of Immigration on Competing Natives' Wages: Evidence from German Reunification. *The Review of Economics and Statistics 102*(1), 79–97.

Rosenfeld, Rachel A., and Heike Trappe (2002). Occupational sex segregation in state socialist and market economies: Levels, patterns, and change in East and West Germany, 1980s and 1998. *Research in Social Stratification and Mobility 19*, 231–267. The Future of Market Transition.

Rosenfeld, Rachel A., Heike Trappe, and Janet C. Gornick (2004). Gender and Work in Germany: Before and After Reunification. *Annual Review of Sociology 30*(1), 103–124.

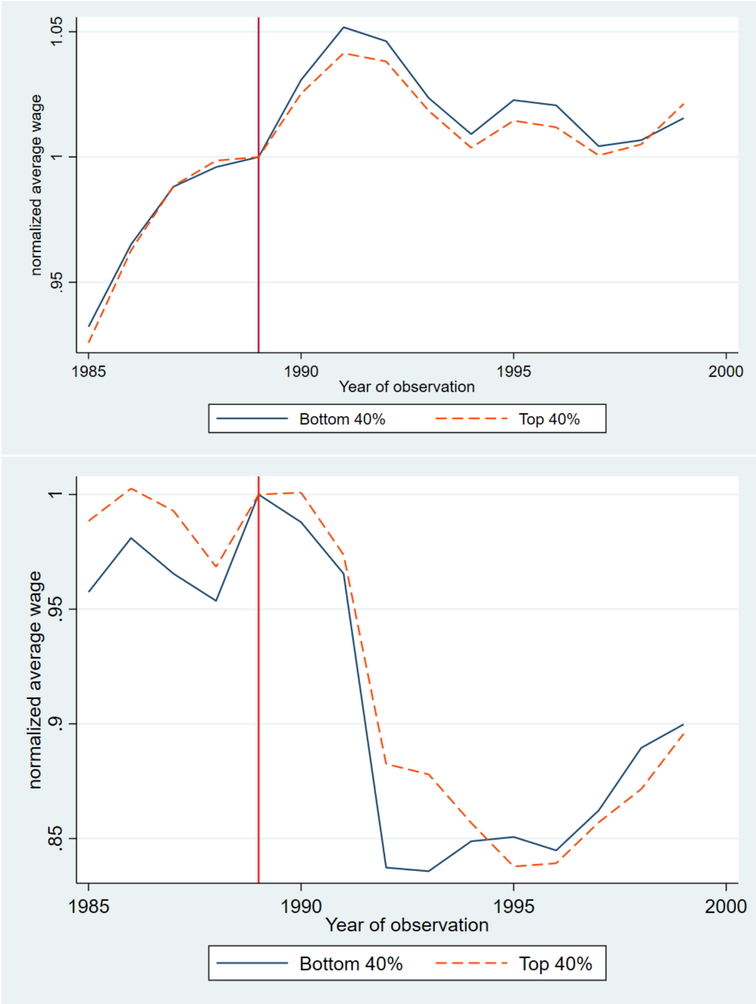
Sanderson, Eleanor, and Frank Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics 190*(2), 212–221.

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

Wegener, Bernd (1985). Gibt es Sozialprestige? *Zeitschrift für Soziologie 14*(3), 209–235.

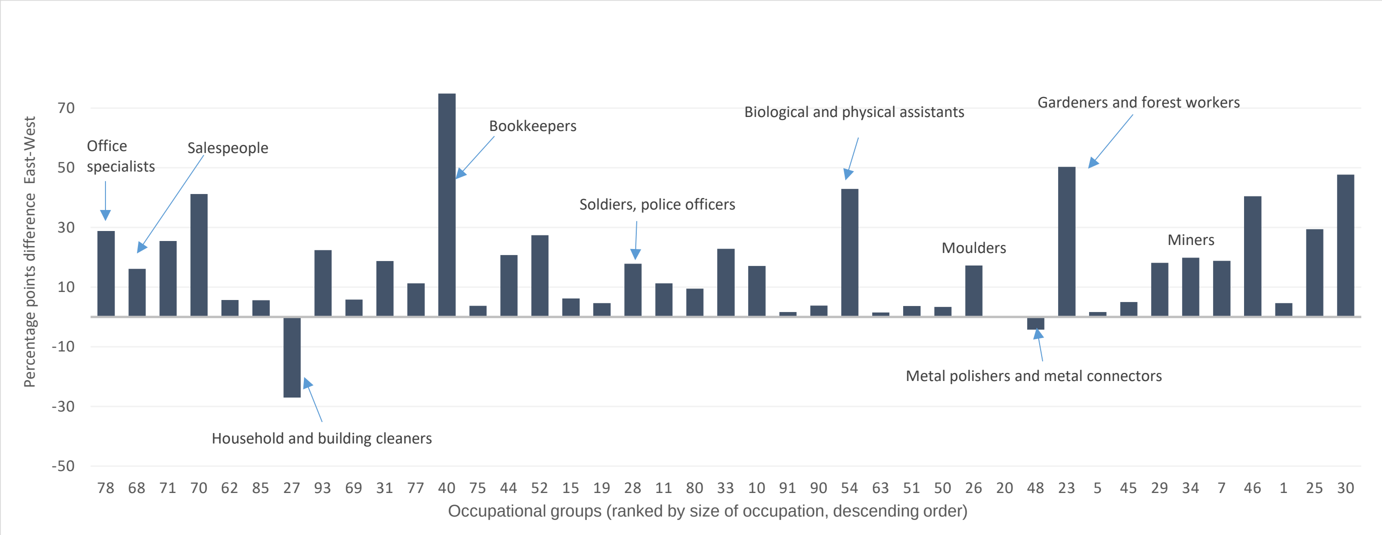
Figures

Figure 1: Evolution of Normalized Average Wages, Men, 1985–1999



Notes: The figure plots the normalized average gross wage for male full-time workers in the years 1985 to 1999 for West Germany (upper figure) and the U.S. (lower figure) for two groups of occupations. The first group comprises the 40 percent of occupations for which East German's female share of training occupations exceeded West German's female share of current occupations the most (Top 40 %). The second group comprises the opposite 40 percent of occupations, for which the female share difference between East and West was the smallest (Bottom 40 %). In the top panel, the sample consists of "native" individuals in the West German labor market aged 25-54, and in the bottom panel, a control group of Americans of the same age in the U.S. labor market. The data for the U.S. sample comes from the Outgoing Rotation Group of the Current Population Survey (for details on the sample construction, see Appendix A).

Figure 2: East-West Difference in Occupational Female Shares

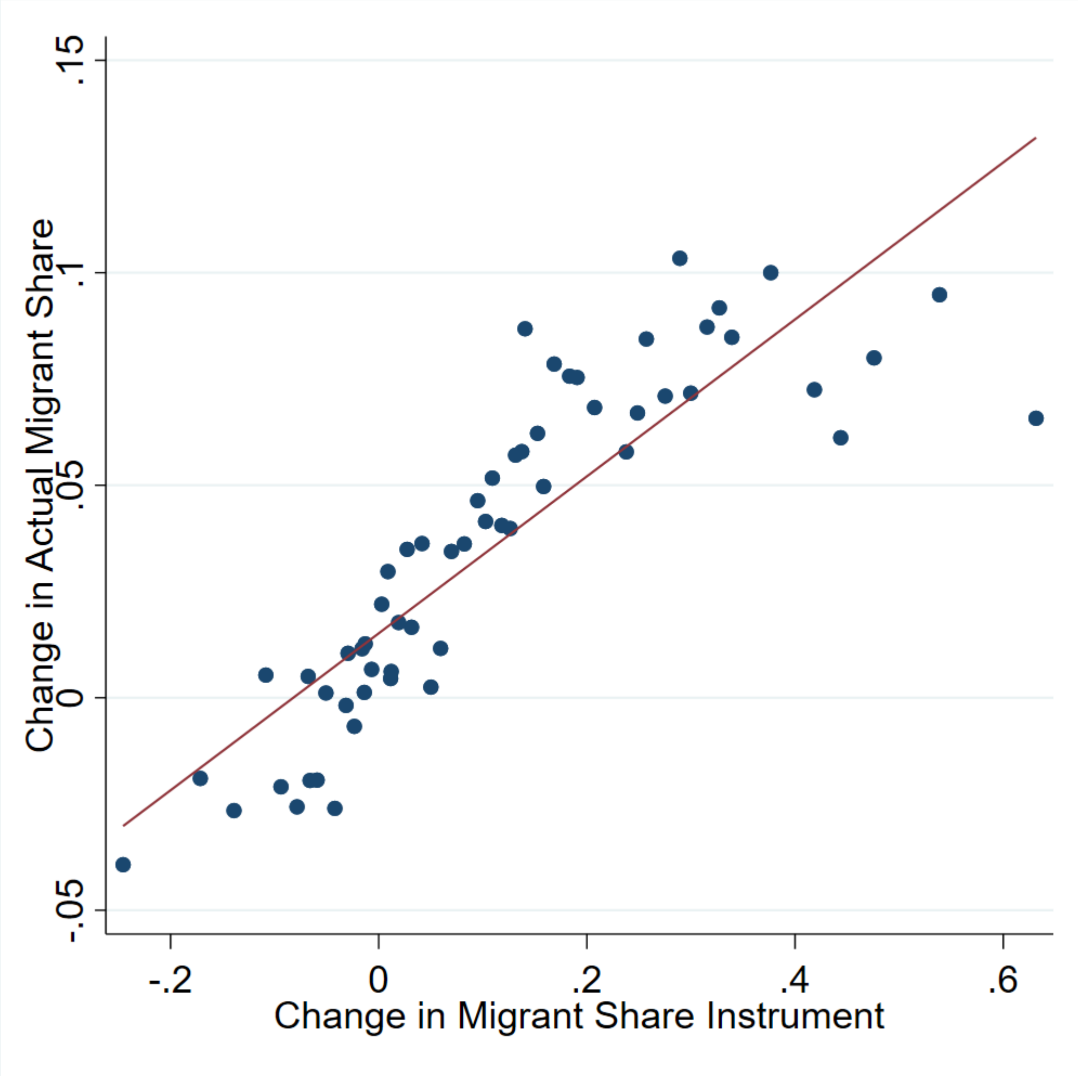


Notes: This figure shows the East-West differences (in percentage points) in female shares in different occupations. Occupations are ordered from left to right, from largest to smallest occupations, according to the SIAB data. East German female shares are measured based on QCS data and relate to female shares in training occupations of East Germans in the 1991/92 and 1998/99 surveys. West-German female shares refer to current occupations and come from the pre-reunification data (1985/86) of the SIAB.

Figure 3: Female Share Instrument:
Graphical Illustration of First Stage

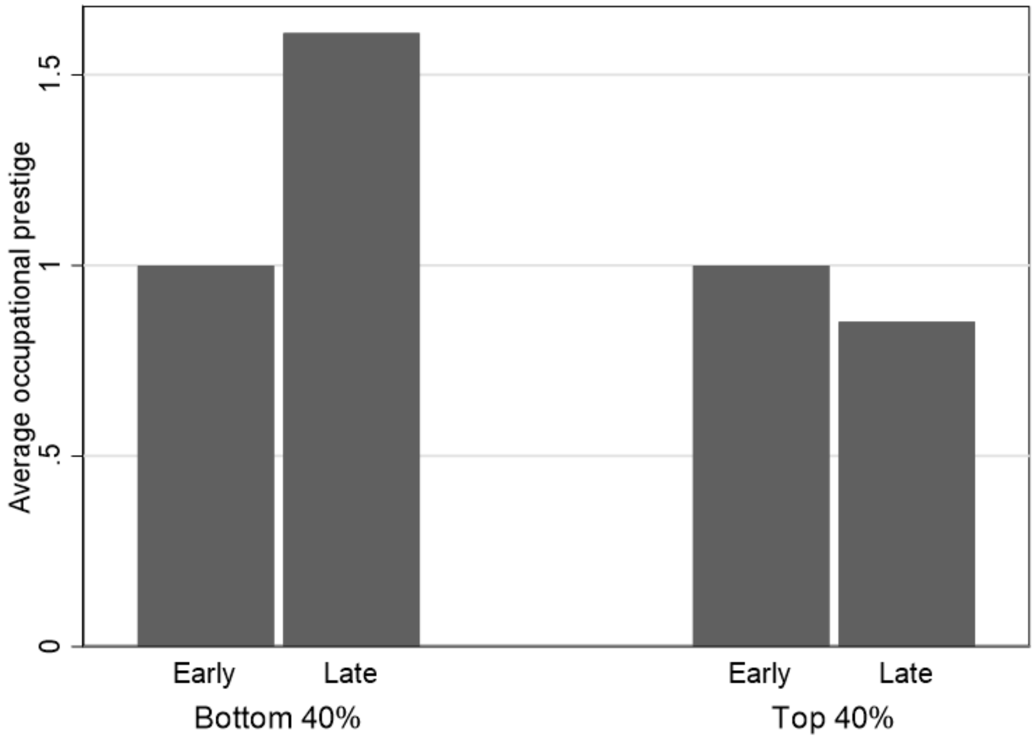
Notes: The graph illustrates the first stage related to the female share instrument. The x-axis shows the average wave-on-wave changes in the female share instrument per age-occupation cell, and the y-axis shows the change in the corresponding actual female share in West Germany. The correlation coefficient is 0.304 and the slope coefficient is 0.215. The sample includes data for 1985/86-1991/92 and 1991/92-1998/99, with a total of 776,669 observations.

Figure 4: Migrant Share Instrument:
Graphical Illustration of First Stage



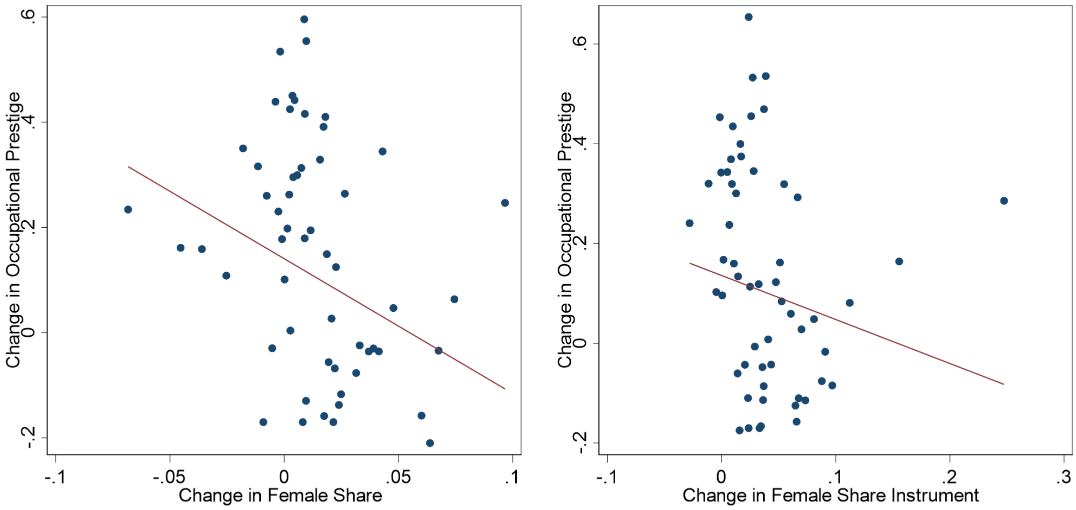
Notes: The graph illustrates the first stage related to the migrant share instrument. The x-axis shows the average wave-on-wave changes in the migrant share instrument per age-occupation cell, and the y-axis shows the change in the corresponding actual migrant share in West Germany. The correlation coefficient is 0.698 and the slope coefficient is 0.185. The sample includes data for 1985/86-1991/92 and 1991/92-1998/99, with a total of 776,669 observations.

Figure 5: Evolution of Occupational Prestige, 1979/1980 to 2017/2018



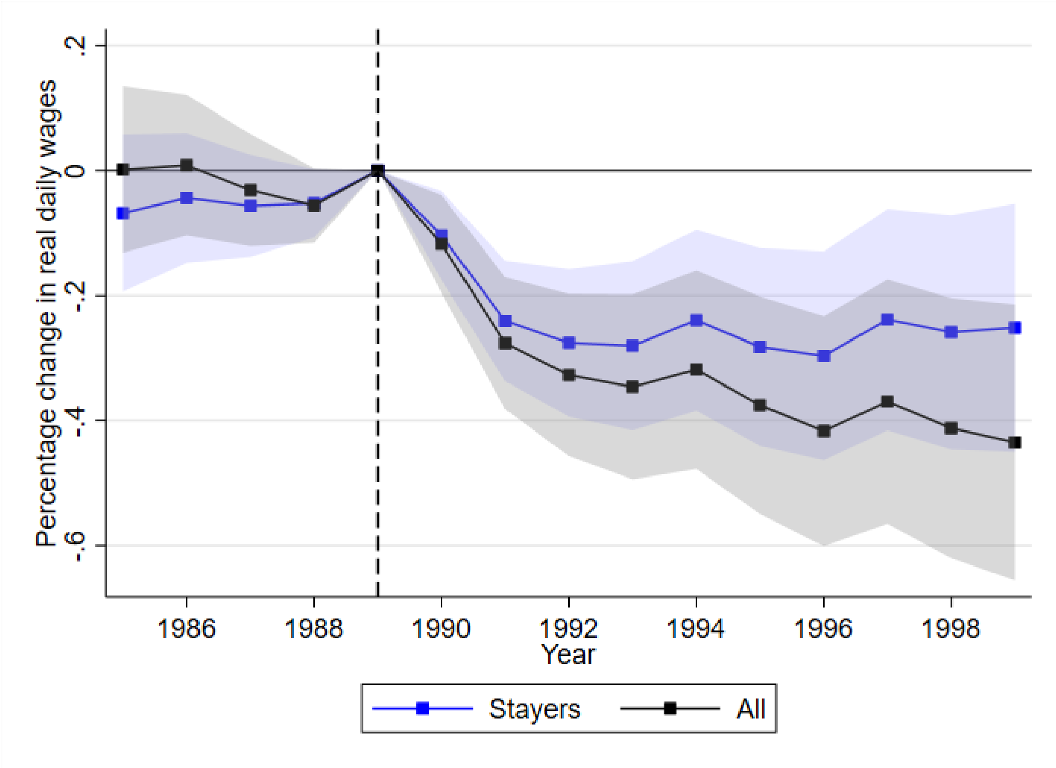
Notes: The figure compares the normalized average prestige scores in the years 1979/80 (pre-reunification period) to 2017/18 (post-reunification period) for two groups of occupations. The first group comprises the 40 percent of occupations for which East German's female share of training occupations exceeded West German's female share of current occupations the most (Top 40 %). The second group comprises the opposite 40 percent of occupations, for which the female share difference between East and West was the smallest (Bottom 40 %). The first group contains 305,811 observations and the second group 478,255, with the sample being the same as for the main analysis. The prestige scores are based on the Magnitude Prestige Scale by Wegener (1985) and the Occupational Prestige Scale (BAS) by Ebner and Rohrbach-Schmidt (2021). The scores for each group are normalized to 1 for the the early period. Further details on the construction of the two prestige scores are provided in Appendix Section A.5.

Figure 6: Changes in the Female Share and Prestige



Notes: The graphs illustrate the correlation between changes in the female share and in occupational prestige over time. The left graph displays the average wave-on-wave changes in the female share per occupation cell on the x-axis, and the change in the prestige score of the corresponding occupation group on the y-axis. The graph on the right displays the average wave-on-wave changes in the female share instrument per occupation cell on the x-axis, with the y-axis corresponding to the same scale as in the left figure. The correlation coefficients are -0.243 and -0.126, respectively, and the slope coefficients for the fitted line are -2.563 and -0.882, respectively. The sample for both figures includes 400,289 individual-level observations for 1985/86 and 1991/92. For the 1985/86 wave, the MPS score is used, and for the 1991/92 wave, the BAS score is used. See Appendix Section A.5 for further details on the construction of the two prestige scores.

Figure 7: Event-Study Approach



Notes: The figure is based on Equation 6. The sample consists of "native" men and women in West German labor market, aged 25-54, who are observed for two consecutive years between 1984 and 1999, respectively. A total of 2,574,018 observations enter the estimation, between 155,906 and 181,147 per year. The coefficients indicate the cumulative effect of the female share shock on log real daily wages, with 1989 as the reference year. The shock is defined as the change in the female share between 1985/86 and 1991/92 in the respective age-occupation cell, and it is instrumented with the potential female share constructed based on QCS data from the respective waves. "All" denotes all observations in the sample, "stayers" denotes those that work in the same occupation in two consecutive sample periods.

Tables

Table 1: Summary Statistics by Year and Overall

	1985/86		1991/92		1998/99		Overall	
	Men	Women	Men	Women	Men	Women	Men	Women
Gross Real Daily Wages								
Mean	108.44	78.39	118.45	88.61	116.01	90.25	114.32	86.19
Median	101.83	77.29	111.19	86.45	108.11	87.47	107.05	83.82
25^{th} percentile	86.50	56.49	94.08	64.70	90.33	64.05	90.00	61.74
75^{th} percentile	124.20	96.37	135.68	108.24	133.95	110.80	131.47	105.67
Female Shares (in Percent) in Age-Occupation-Time Cell								
Mean	23.16	63.98	25.61	64.90	26.07	63.49	24.93	64.14
Median	12.07	71.01	15.59	72.68	14.45	70.67	14.01	71.01
25^{th} percentile	2.77	54.03	4.05	57.10	4.16	53.79	3.73	54.03
75^{th} percentile	41.78	76.50	44.65	78.24	45.75	75.23	44.65	76.50
Number of Employees in Age-Occupation-Time Cell								
Mean	3,728	8,030	4,636	10,196	4,763	10,063	4,373	9,518
Median	2,759	7,548	3,291	9,080	3,304	8,118	3,079	8,115
25^{th} percentile	1,408	2,412	1,681	3,282	1,456	3,546	1,467	2,984
75^{th} percentile	4,426	14,366	5,380	17,398	5,599	16,685	5,114	16,663
East German Migrant Shares (in Percent) in Age-Occupation-Time Cell								
Mean	0.00	0.00	7.42	6.05	7.45	5.93	4.94	4.24
Median	0.00	0.00	7.14	5.34	6.54	5.66	4.77	4.48
25^{th} percentile	0.00	0.00	4.70	4.45	4.97	3.94	0.00	0.00
75^{th} percentile	0.00	0.00	9.38	7.49	9.91	7.10	8.46	6.23
N	259,413	107,735	266,733	133,556	248,657	127,723	774,803	369,014

Notes: The table shows summary statistics for our main sample, including both men and women. The sample includes only full-time working individuals between 25 and 54 years old with a medium level of education. We show the figures for the occupational female shares and migrant shares in percent in this table, but we use these variables divided by 100 in the estimations.

Table 2: FEMINIZATION OF OCCUPATIONS AND THE EFFECT ON WAGES: RESULTS FROM OLS AND IV REGRESSIONS

Dependent variable: Log real daily wages of West Germans (w_{iaot})						
	OLS	IV	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Female share (f_{aot})	-0.498***	-0.691***	-0.867***	-0.559***	-0.573***	-0.832***
	(0.070)	(0.090)	(0.168)	(0.073)	(0.079)	(0.314)
$f_{aot} \times \mathbb{1}$(District w/ Increase)			-0.045***			
			(0.010)			
$f_{aot} \times \mathbb{1}$(Inflow)				-0.137***	-0.134***	
				(0.014)	(0.014)	
$f_{aot} \times \mathbb{1}$(Outflow)				-0.007	-0.009	
				(0.008)	(0.008)	
$f_{aot} \times \mathbb{1}$(Older than 39 years)					0.532***	
					(0.168)	
$f_{aot} \times \mathbb{1}$(0.25-0.5 female share)						0.354
						(0.344)
$f_{aot} \times \mathbb{1}$(0.5-0.75 female share)						0.318
						(0.356)
$f_{aot} \times \mathbb{1}(\geq$ 0.75 female share)						0.289
						(0.397)
Migrant share (m_{aot})	-0.220***	-1.190***	-1.270***	-0.827**	-0.821**	-1.051***
	(0.067)	(0.409)	(0.441)	(0.339)	(0.327)	(0.376)

Notes: Pooled cross-sectional data of full-time workers in years 1985, 1986, 1991, 1992, 1998 and 1999. The number of observations is 774,540 in Column (3) and 1,143,817 in all other Columns. As indicated in Equation 1, all specifications include a gender dummy and age-occupation-time (-gender) fixed effects. The regression corresponding to Column (3) additionally contains a dummy variable that takes one the value one for districts which report an increase in the female share between two waves. The specifications in Columns (4) and (5) additionally include baseline dummies associated with the respective interaction terms (inflow, outflow, older than 39). Standard errors (in parentheses) are clustered at the age-occupation-time level. Detailed first-stage and reduced-form regression results for the specification in Column (2) are shown in Table 3. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: FEMINIZATION OF OCCUPATIONS AND THE EFFECT ON WAGES: RESULTS FROM FIRST STAGE AND REDUCED FORM REGRESSIONS

	First Stages		Reduced Form
	(1)	(2)	(3)
Dependent variable:	Female Share	Migrant Share	Log Real Daily Wages
Female share instrument (f_{aot}^{IV})	0.507***	-0.012	-0.336***
	(0.042)	(0.014)	(0.044)
Migrant share instrument (m_{aot}^{IV})	-0.088***	0.057***	-0.007
	(0.018)	(0.011)	(0.024)
F-test:	71.21	13.09	31.19
SW F-test:	129.47	26.97	

Notes: Pooled cross-sectional data of full-time workers in years 1985, 1986, 1991, 1992, 1998 and 1999. The number of observations is 1,143,817 in all specifications. Columns (1) and (2) show first-stage results, and column (3) shows reduced form results. All specifications include controls identical to the relevant second-stage specification shown in Table 2, column (2). We report Sanderson and Windmeijer (SW) F-test statistics. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Feminization of Occupations and the Effect on Tasks (2SLS, QCS Data)

Dependent variable: Worker's share of tasks in...task category					
	(1)	(2)	(3)	(4)	(5)
	Analytical	Interactive	Routine Cognitive	Routine Manual	Non-routine Manual
Female share (f_{aot})	-0.119*	-0.118*	0.105	0.099	-0.033
	(0.053)	(0.047)	(0.076)	(0.098)	(0.083)
Migrant share (m_{aot})	-0.151	-0.068	-0.086	0.120	0.029
	(0.105)	(0.102)	(0.177)	(0.216)	(0.169)

Notes: Pooled cross-sectional data of full-time workers from the QCS data. The number of observations is 16,012 in all specifications. As indicated in Equation 1, all specifications include a gender dummy and age-occupation-time (-gender) fixed effects. Full first-stage results are displayed in Table C.5. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Feminization of Occupations and the Effect on Wages: Results from IV Regressions

Dependent variable: Log real daily wages of West Germans (w_{iaote})						
	IV (1)	IV (2)	IV (3)	IV (4)	IV (5)	IV (6)
Female share (f_{aot})	-0.536***	-1.092***	-0.684***	-0.690***	-0.582***	-0.713***
	(0.098)	(0.113)	(0.088)	(0.088)	(0.094)	(0.091)
f_{aot} * $\mathbb{1}$(Small establishment)	-0.239***					
	(0.018)					
f_{aot} * $\mathbb{1}$(No entry restriction)		0.392***				
		(0.031)				
f_{aot} * $\mathbb{1}$(Service sector)			0.006			
			(0.008)			
f_{aot} * $\mathbb{1}$(Low-Skill Services)				-0.040***		
				(0.010)		
f_{aot} * $\mathbb{1}$(High-Skill Services)				0.016		
				(0.012)		
f_{aot} * $\mathbb{1}$(Medium AKM)					-0.021*	
					(0.011)	
f_{aot} * $\mathbb{1}$(Lowest AKM)					-0.190***	
					(0.015)	
f_{aot} * $\mathbb{1}$(Medium HHI)						0.022***
						(0.007)
f_{aot} * $\mathbb{1}$(High HHI)						0.039**
						(0.018)
Migrant share (m_{aot})	-1.255***	-1.370***	-1.035***	-1.070***	-1.119***	-1.185***
	(0.409)	(0.443)	(0.388)	(0.385)	(0.396)	(0.411)

Notes: Pooled cross-sectional data of full-time workers in years 1985, 1986, 1991, 1992, 1998 and 1999. Each regression is based on more than 1 million observations. As indicated in Equation 1, all specifications include age-occupation-time (-gender) fixed effects. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Gendered Effect on Wages:
Results from OLS and IV Regressions

Dependent variable: Log real daily wages of West Germans (w_{iaot})				
	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)
Female share (f_{aot})	-0.582***	-0.802***	-0.682***	-0.736***
	(0.057)	(0.101)	(0.075)	(0.107)
Female share (f_{aot}) × female	0.246***	0.323**	0.610***	0.354***
	(0.072)	(0.133)	(0.105)	(0.103)
Migrant share (m_{aot})	-0.215**	-1.143***	-1.080***	-0.987***
	(0.084)	(0.405)	(0.362)	(0.362)
Individual fixed effects	No	No	Yes	No
Firm fixed effects	No	No	No	Yes

Notes: Pooled cross-sectional data of full-time workers in years 1985, 1986, 1991, 1992, 1998 and 1999. The number of observations is 1,143,817 in all specifications. As indicated in Equation 1, all specifications include age-occupation-time (-gender) fixed effects. The specification in Column (3) additionally includes individual fixed-effects. Standard errors (in parentheses) are clustered at the age-occupation-time level. Detailed first stage and reduced form regression results are shown in Table C.3. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix

A Data

A.1 SIAB: Data preparation

The Integrated Labor Market Biographies (SIAB) sample is our main analysis sample. Specifically, we base our analysis on the most recent and weakly anonymized version, "SIAB File 7519". To prepare the SIAB data for our purposes, we follow Dauth and Eppelsheimer (2020). We follow Fitzenberger et al. (2006) when it comes to identifying the education level of the observations. We use all available information from both employment and unemployment (from the "Leistungsempfängerhistorik", LeH) to identify the East or West German origin of workers based on their entire social security history in the data (for details, see Section A.4 below).

Regarding the transformation from spell structure to panel structure, we select the spell with the longest employment duration (and the highest wage in case of equal duration) as the main spell at any given point in time. If a person has several main spells in one year due to changing employers, we choose the one with the longest duration and highest wage to get one observation per person and year.[36] We modify our procedure - compared to Dauth and Eppelsheimer (2020), who always choose the spell that covers June 30 each year - to increase the representativeness of the selected spell for the entire year. Regarding the wage imputation, we also make a small adaption to the algorithm. Instead of running Tobit-wage-regression for every year, education group, and East/West separately, we distinguish between year, East/West, and gender groups for the wage estimation and use the education categories as explaining variables in the regressions.

We also restrict the SIAB data to fit the QCS data that is used to construct the instruments. First, for our cross-sectional analyses, we restrict the observations to the years 1985, 1986, 1991, 1992, 1998, and 1999 so the SIAB data fit the QCS waves of 1985/86, 1991/92, and 1998/99. Since the SIAB reports yearly data, while the data collection of the QCS overlaps with the next year, we exploit the information of two years for each QCS wave. For the event-study analyses, we keep 1980 to 1999 and work with fixed instruments that we apply to all survey years equally. We restrict our sample to employees between 25 and 54 years old who pay social insurance contributions and are not considered marginal workers. Observations are grouped based on six age groups, the three survey waves, and 42 occupational categories for which the QCS instruments are available (see Section A.3 below). Based on this data, we calculate female shares, employment, and East German shares in each occupation-age-wave cell in West Germany. To construct the instruments, we combine the data on potential East German labor supply from the QCS with numbers on West German employment before reunification from the SIAB, as described in Section 4.2.

For the estimation sample, we drop all workers currently employed in East Germany and those we identify as East Germans based on their employment history (see below) to capture only the wage effects on West German natives. We further restrict the analysis to people with a medium level of education (i.e., an apprenticeship) and complete occu-

[36]In case there is no unique spell with the highest wage and the longest duration, we pick the one with the highest income.

pation and firm information. Finally, we construct a sample used consistently across all specifications by dropping workers observed only once in the data as they will provide no information in specifications with individual fixed effects. The same applies to workers at firms where we only have one observation, which would be dropped in any regression with establishment fixed effects. Furthermore, we restrict the sample to men and women employed full-time for our main analysis. The summary statistics in Table 1 as well as the regression results in Tables 2 and 3 are all based on this sample. Results for a sample that includes part-time workers can be found in the Appendix in Table C.2.

A.2 The Current Population Survey (CPS)

To conduct a graphic placebo test to see whether we find diverging trends for different groups of occupations regardless of the reunification shock, we replicate our analysis with data from the Current Population Survey (CPS) in the bottom panel of Figure 1. We construct a sample comparable to that from the SIAB, using the monthly outgoing rotation group survey and restricting our analysis to men aged 25-54. Again, we limit the sample to individuals with a medium education level, which we define as having completed high school but not college. We use average weekly earnings (conditional on working at least one week) to compute averages by occupation group using earning weights provided by the CPS. We group the individuals into occupations with a low and high (placebo) female share shock, respectively, by assigning the occupation classifications from the CPS to the occupation groupings from the SIAB/QCS data. The procedure is described in more detail below. Our final sample consists of 382,393 individuals, of which 304,599 (80%) can be assigned to a female share shock category and enter the graph.

A.3 Occupational classification

The QCS and SIAB data use the same classification system for occupations, the Classification of Occupations (KldB) 1988 by the Federal Employment Agency of Germany. Our analysis is on the occupation-age-time level, so we need to aggregate them to have sufficient non-empty cells. In their original version, the occupations are classified on the 3-digit level. In the analyses, we use a hybrid version in which some occupations are still on the 3-digit level, and some are on the 2-digit level, depending on the number of observations in the occupation-age cells. In addition, for constructing the instrument, we only consider occupations where vocational training is the usual form of education ("anerkannte Ausbildungsberufe"). The overall procedure results in a final set of 42 occupations.

CPS data treatment and occupation classification: For the bottom panel of Figure 1, we manually match the 389 categories provided by the OCC1990 variable in the IPUMS version of the CPS (Flood et al., 2022) with the 42 occupational categories from above, using the descriptions of the underlying 3-digit categories as additional guidance. Two independently obtained classifications were identical for 283 out of 389 occupations (73% of the occupations and 67% of the data), representing 80% of the original sample (44.7% in the "bottom 40%" category and 40.5% in the "top 40%" category).

A.4 Identification of East Germans

Unlike the QCS, the SIAB does not provide an indicator that directly identifies East and West Germans. To calculate the share of migrants in each cell and to remove the East Germans from the estimation sample, we construct an identifier for East Germans based on their location and history within the social security data. Hereby, the key element to identify the place of origin of an individual is the location of the first observation that is available in the data. We start with the full set of spells for employment (BeH) and unemployment (LeH).

As a first step, we classify the current location of an observation as East or West German based on the information available on the firm's location ("ao_bula", from the firm data set BHP). For the 8,108,075 (out of 50,561,644 total) unemployment spells, this information is naturally not available (information on the location of the regional unemployment agency ("wo_rd") and sometimes on the place of residence ("wo_bula") is given instead and will be used in a later step). Additionally, there are 6,806 employment spells for which the firm identifier is missing and can not be assigned a firm location. Lastly, for 1,043,305 observations, the firm identifier is available, but it cannot be matched to the firm data. This can happen for two reasons: First, the firm data is reported on June 30 each year, and firms that either closed down before the deadline or were opened after it are not recorded. Second, for East Germany, the firm reporting structure was only established in 1992, one year after the East German workers were added to the social security records. Therefore, no location information is available for all observations in East Germany in 1991. Additionally, because the reporting system was set up only slowly, there was an unusually high amount of firms with missing information in 1992. We address these timing and incomplete reporting issues by filling in the location information from an adjacent year if the establishment always reports the same location or the missing entry is between two valid identical firm location observations. This adds locational information for 698,484 observations, leaving us with 8,459,702 out of 50,561,644 million observations without a firm location and 351,798 for which no kind of locational info (place of work, place of residence, or place of unemployment agency) whatsoever is present. We convert this information into East and West dummies, with Berlin being classified as West Germany until 1991.[37] We assign it to neither category afterward. With this procedure, we classified 5,085,311 observations as East German (10%), 37,152,172 as West German (73%), and 8,324,161 as unassigned (16%).

We fill in missing information where appropriate to further increase the number of observations. Suppose all valid locations of a person are in East Germany. In that case, we assume that this person has spent their entire life in East Germany, and any unclassified observations are also in East Germany. For individuals with a fully West German work history, we apply the same rule but exclude persons whose first spell happened between 1990 and 1993 to avoid misclassifying people who moved during the early reunification phase. This step adds information for 5,957,427 observations. Next, we replace missings between two valid identical location classifications, adding another 1,5490,68 observations. Overall, 817,666 observations (1.6%) remain unclassified. All analyses are based on observations unambiguously located in West Germany, dropping observations from

[37]Before establishing the firm reporting system in 1992, observations from East Berlin could not have entered the SIAB data with valid location information. Therefore, any observation with the location information "Berlin" in 1991 or earlier must be located in West Berlin.

East German establishments and unclassified observations.

A.5 Construction of Occupational Prestige Scores

For the analysis of the development of occupational prestige over time, we adopt prestige scores from two different sources. For the pre-reunification period, we rely on the Magnitude Prestige Score constructed by Wegener (1985). This score aims to capture the value and recognition that an occupation and its members receives. Prestige is expected to correlate with other measures of social status, such as income and education, but adds additional elements of the "reputation" or "'recognition" of an occupation. The prestige score presented by Wegener (1985) is based on surveys conducted in 1979 and 1980 in West Germany.

We use the version adopted to the KldB1992 three-digit level by Frietsch and Wirth (2001) and transfer the scores to the slightly different KldB1988 codes by using the correspondence table implied in the QCS 1991/92 data (as used above for the construction of the instrument). For each observation in the QCS, both a KldB1992 and a KldB1988 occupation are listed. In case of a one-to-one match, we can directly assign the prestige score provided by Frietsch and Wirth (2001) to the KldB1988 three-digit occupation. In case there are multiple entries that correspond to the same KldB1988 code, we form a weighted average of the corresponding scores, using the number of QCS observations assigned to the respective KldB1992 categories as weights.

For the post-reunification period, we use the occupational prestige scores by Ebner and Rohrbach-Schmidt (2021). These prestige scores were constructed in an attempt to update the Wegener (1985) scores to reflect changed occupational structures and valuations. Surveys asking about the prestige of 402 different occupations were conducted in the German population in 2017 and 2018.[38] The survey responses are then used to construct scores at various KldB2010 levels. We use the scores at the three-digit level constructed by taking averages over the underlying groups (BAS-3-V1) and assign the scores directly to our sample observations, using the fact that in the SIAB data, both the KldB1988 and KldB2010 occupations are reported for each entry.

Next, for both scores, we compute values at the occupation level corresponding to the one we use in the analyses by taking weighted averages. The methodology of the two scores varies slightly but is sufficiently similar to warrant comparisons over time. Ebner and Rohrbach-Schmidt (2021) conduct a validity check of their methodology in which they compare their score with the 1979/1980 score, finding a correlation of 0.63 at the 3-digit level. They conclude that the scores are comparable over time and that the variations in scores reflect both methodological differences (particularly the conversion between different occupational classifications) and trends in prestige over time. In our sample, we find a correlation of 0.34 at the 3-digit level and a correlation of 0.28 at the occupation group level.

[38] Note that while the original MPS scores covered West Germany only, the new scores are based on data from all of Germany.

A.6 QCS sample for task-based analyses

For the exploration of changes in the task structure, the same QCS sample as for the instrumental variable is used. Task information is available at the individual level. The five task categories follow the ones developed in Spitz-Oener (2006) and are based on a set of survey responses about an individual's tasks performed at the workplace. The final regression sample is constructed using all West Germans working in recognized apprenticeship occupations ("anerkannte Ausbildungsberufe"), with medium education, and excluding civil servants and certain occupations and industries which could also not be used for the construction of the instrument. The female share and migrant share instruments are constructed as before, while the actual female share and migrant share are calculated based on QCS data only. The final sample contains 16,010 observations from both men and women working in full-time jobs (defined as working at least 20 hours per week).

A.7 Definition of additional variables used in the heterogeneity analyses in Table 5

Manufacturing vs Services. We use the 3-digit industry codes provided in the Establishment History Panel to classify the establishments into 15 industries following the classification provided by the Statistical Office.[39] We then further aggregate the industries into manufacturing (categories 2 to 4) and services (categories 5 to 15). In a further classification, the services sector is split up into high- and low-skill services (with 5, 6, 7, 11, and 14 falling into low- and 8, 9, 10, 12, 13, and 14 falling into high-skill services) following the methodology by Bárány and Siegel (2018).

Abowd, Kramarz, and Margolis (AKM, 1999) establishment fixed effects. We use the AKM fixed effects provided by the IAB, as described in Lochner et al. (2023). Based on the methodology by Card et al. (2013), imputed wages are regressed on time-varying characteristics and individual and firm fixed effects. The individual fixed effects can be seen as a proxy for unobserved ability (constant over time and independent of the employer), while the firm fixed effects represent the average pay premium or discount at the establishment, which could be caused, for example, by differences in productivity or market power. In our context, we are interested in whether firms that pay above- or below-average wages relative to their competitors react differently to an inflow of females into an occupation. We use the raw values calculated for West German establishments from 1985-1991 to group establishments into AKM terciles based on estimated AKM establishment fixed effects.

Labor market concentration. We compute flow-adjusted employment Herfindahl-Hirschman indices (HHI) adapting the methodology of Arnold (2021). The approach provides measures of labor market concentration on the occupation-region-time level (see Equation A1, in which o indicates occupations, r indicates regions, and t refers to time).

[39]The industries are (1) Missings (reference category), (1) Agriculture, mining, and quarrying, (2) Manufacturing, (3) Energy and other infrastructure, (4) Building industry, (5) Sale and Maintenance, (6) Transport and Logistics, (7) Hospitality, (8) Information, (9) Finance, (10) Land and housing rentals, (11) Other services (12) Public services, (13) Education and health, (14) Private household services, and (15) Extraterritorial organizations. See `https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/gliederung-klassifikation-wz-3100130089004.pdf?__blob=publicationFile`.

It measures local concentration and includes job-to-job mobility patterns to allow for substitutability of workers across establishments.

As shown in Equation A1, we compute the HHI from the sum of flows-adjusted squared employment shares, $\tilde{s}_{jort}$, over all establishments in a particular region-occupation-year.

Equation A2 defines establishment j's flows-adjusted employment share. The numerator is the weighted sum of establishment j's employment level l_{jkrt} over all occupations in a particular region-year. The denominator is the weighted sum of total employment L_{krt} over all occupations in a particular region-year. Equation A3 defines the weights that enter equation A2. The term $E\left[\frac{L_k}{L_o}\right]$ expresses the expected relative size of total employment in the respective occupation across labor market regions. It is worth noting that the weight on occupation k increases if workers in occupation o are likely to transition to occupation k, especially when total employment of occupation k is small. We compute the weighting factors from transition rates between occupations in a pooled sample covering 1993 to 2010. To apply the HHI indices to our analysis, we focus on the values from the year 1988. While the original index is calculated on the three-digit occupation level, we aggregate it (weighted by total employment in the cell) to our 42 occupation groups. We then group the occupation-region cells into low-, middle-, and high-concentration terciles.

$$ArnoldHHI_{ort} = \sum_{j=1}^{N} \tilde{s}^2_{jort} \tag{A1}$$

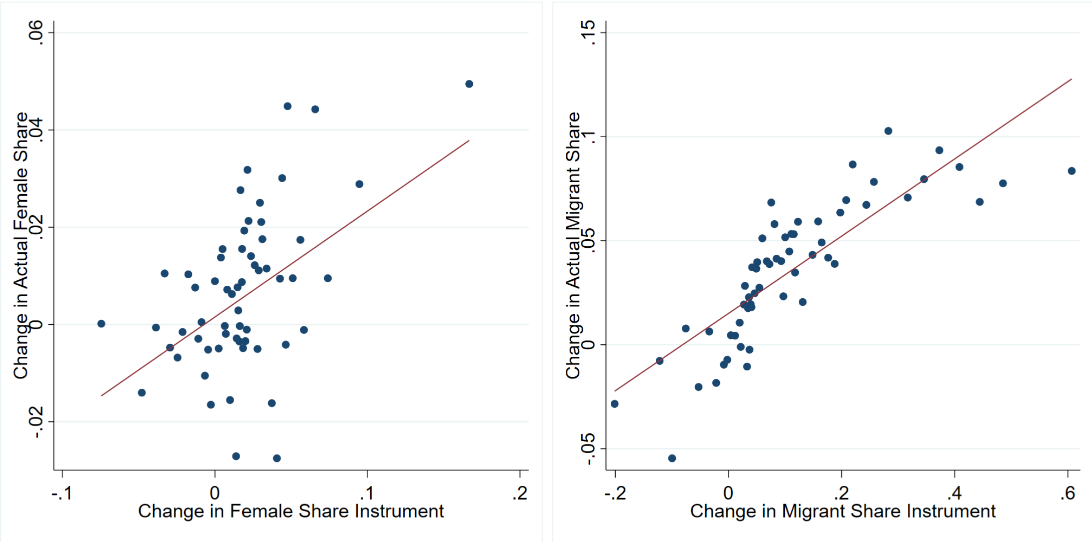
$$\tilde{s}^2_{jort} = \frac{\sum_{k \in r} \alpha_{o \to k} l_{jkrt}}{\sum_{k \in r} \alpha_{o \to k} L_{krt}} \tag{A2}$$

$$\alpha_{o \to k} = \frac{P(k|o)}{P(o|o)} \frac{1}{E\left[\frac{L_k}{L_o}\right]} \tag{A3}$$

For each observation, we record the location of the employer in 1988 (or of another year if 1988 is not available and the individual works in the same commuting zone in all observed years) and assign it to a commuting zone based on the classification by the Federal Statistical Office. We then merge the HHI indices and terciles to the SIAB data, with a match rate of 95.6%.

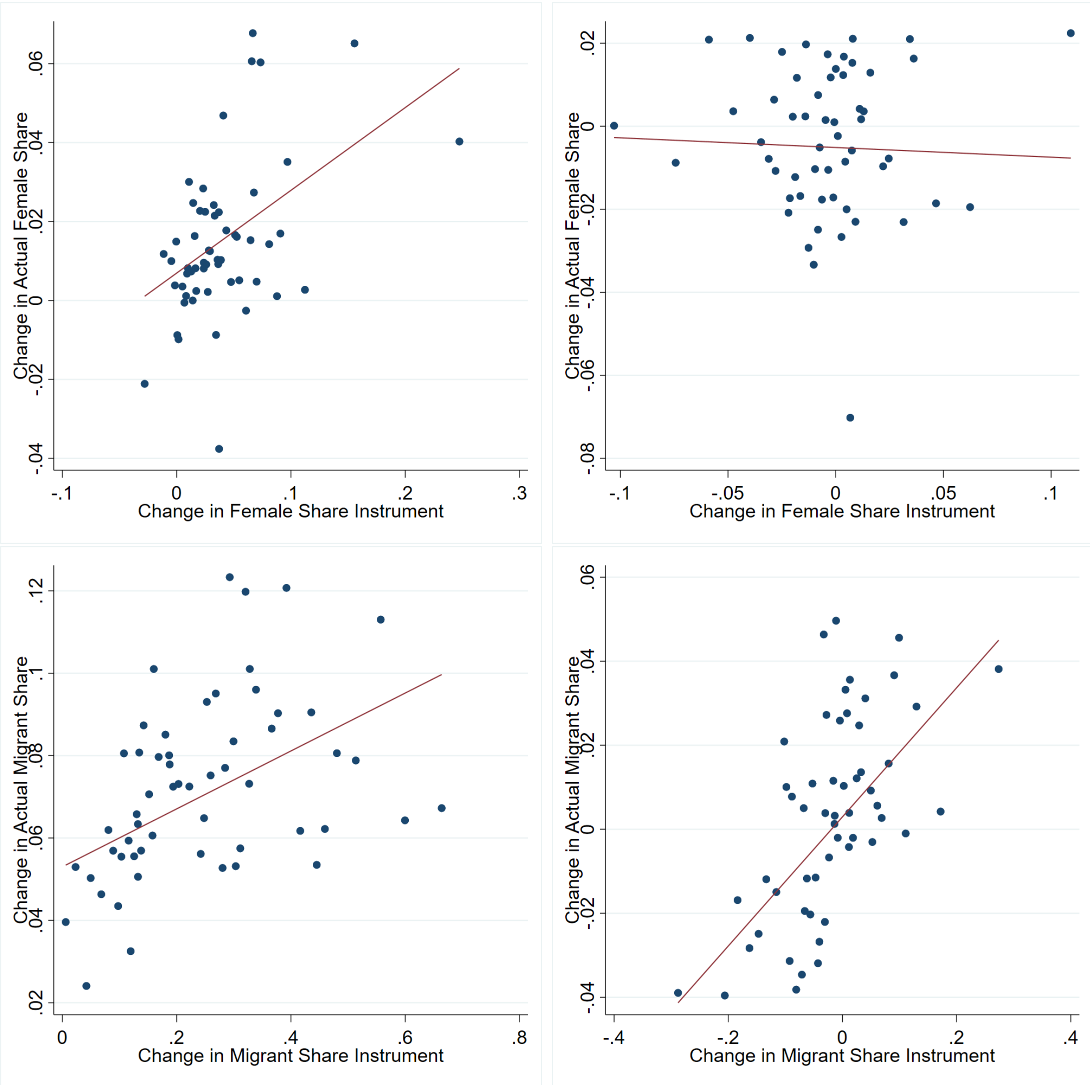
B Figures

Figure B.1: Graphical Illustration of First Stage: Controlling for the Other Endogenous Variable



Notes: The graph illustrates the first stage related to the female share instrument (left panel) and migrant share instrument (right panel), controlling for the changes in the respective other endogenous variable. The x-axis shows the average wave-on-wave changes in the instrument per age-occupation cell, and the y-axis shows the change in the corresponding endogenous variable, net of changes in the respective other instrument. The slope coefficients are 0.218 for the female share regression and 0.186 for the migrant share regression. The sample size is 776,669.

Figure B.2: Instruments:
Graphical Illustration of First Stage, Separately by Wave



Notes: The graph illustrates the first stage related to the female share instrument (top panel) and migrant share instrument (bottom panel). The change from 1985/86 to 1991/92 is shown on the left side, the change from 1991/92 to 1998/99 on the right side. The x-axis shows the average wave-on-wave changes in the respective instrument per age-occupation cell, the y-axis shows the change in the corresponding actual female/migrant share in West Germany. For the female share, the correlation coefficients are 0.316 (1985/86 to 1991/92) and -0.024 (1991/92 to 1998/99), and the slope coefficients are 0.210 and -0.023. For the migrant share, the correlation coefficients are 0.340 and 0.419, respectively, and the slope coefficients are 0.070 and 0.154. The sample sizes are 400,289 (1985/86 to 1991/92) and 376,380 (1991/92 to 1998/99).

C Tables

Table C.1: Summary Statistics by Year and Overall: Instruments

	1985/86		1991/92		1998/99		Overall	
	Men	Women	Men	Women	Men	Women	Men	Women
Female Shares (in Percent) in Age-Occupation-Time Cell								
Mean	23.16	63.98	28.11	66.89	28.08	64.93	26.44	65.36
Median	12.07	71.01	18.36	74.34	17.86	73.51	15.28	73.51
25^{th} percentile	2.77	54.03	5.52	59.36	5.27	58.01	4.81	56.71
75^{th} percentile	41.78	76.50	49.02	78.08	47.78	76.93	46.70	77.50
East German Migrant Share Instrument (in Percent) in Age-Occupation-Time Cell								
Mean	0.00	0.00	26.88	17.02	22.52	13.07	16.48	10.68
Median	0.00	0.00	26.41	13.79	19.13	13.75	12.30	9.76
25^{th} percentile	0.00	0.00	14.20	9.76	8.02	3.02	0.00	0.00
75^{th} percentile	0.00	0.00	37.02	22.83	34.55	17.80	29.93	15.84
N	259,413	107,735	266,733	133,556	248,657	127,723	774,803	369,014

Notes: The table shows summary statistics for the full sample. For both genders, the samples include individuals between 25 and 54 years old with a medium level of education. We show the figures for the female share and migrant share instruments in percent in this table, but we use theses variables divided by 100 in the estimations.

Table C.2: FEMINIZATION OF OCCUPATIONS AND THE EFFECT ON WAGES: ROBUSTNESS IV REGRESSIONS

Dependent variable: Log real daily wages of West Germans (w_{iaote})				
	excl. 98/99 (1)	excl. 98/99 (2)	Full+Part Time (3)	Full+Part Time (4)
Female share (f_{aot})	-0.613***	-0.732***	-0.565***	-0.790***
	(0.080)	(0.090)	(0.081)	(0.102)
Female share (f_{aot}) × female		0.352***		0.543***
		(0.122)		(0.124)
Migrant share (m_{aot})	0.134	0.129	-1.209**	-1.107***
	(0.312)	(0.310)	(0.408)	(0.396)
N	767,437	767,437	1,306,261	1,306,261

Notes: Pooled cross-sectional data of full-time workers in 1985, 1986, 1991, 1992, 1998, and 1999. In columns (1) and (2), data from 1998 and 1999 are excluded; in columns (3) and (4), part-time workers are added to the sample. All specifications include age-occupation-time(-gender) fixed effects, and columns (2) and (4) additionally include gender dummies. Columns (3) and (4) additionally include part-time dummies. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C.3: Feminization of Occupations and the Effect on Wages: Results from First Stage and Reduced Form Regressions – With Female Interaction

	First Stages			Reduced Form
	(1)	(2)	(3)	(4)
Dependent variable:	Female Share	Female Share × Female	Migrant Share	Log Real Daily Wages
Female share instrument (f_{aot}^{IV})	0.486***	0.024***	-0.013	-0.367***
	(0.043)	(0.009)	(0.016)	(0.050)
Female share instrument (f_{aot}^{IV}) × female	0.066***	0.479***	0.004	0.097*
	(0.024)	(0.045)	(0.013)	(0.058)
Migrant share instrument (m_{aot}^{IV})	-0.088***	-0.038***	0.057***	-0.006
	(0.018)	(0.006)	(0.011)	(0.024)
F-test:	52.46	52.48	8.94	20.70
SW F-test:	118.08	116.05	29.12	

Notes: Pooled cross-sectional data of full-time workers in years 1985, 1986, 1991, 1992, 1998 and 1999. The number of observations is 1,143,817 in all specifications. Columns (1), (2), and (3) show first-stage results, and column (4) shows reduced form results. All specifications include controls identical to the relevant second stage specification shown in Table 6, column (2). We report Sanderson and Windmeijer (SW) F-test statistics. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C.4: Feminization of Occupations and the Effect on Tasks (OLS, QCS data)

Dependent variable: Worker's share of tasks in...task category					
	(1)	(2)	(3)	(4)	(5)
	Analytical	Interactive	Routine Cognitive	Routine Manual	Non-routine Manual
Female share (f_{aot})	-0.024	-0.022	0.044	0.017	-0.043
	(0.021)	(0.022)	(0.033)	(0.040)	(0.034)
Migrant share (m_{aot})	0.018	-0.004	-0.063	0.003	0.050
	(0.025)	(0.030)	(0.046)	(0.048)	(0.051)

Notes: Pooled cross-sectional data of full-time workers from the QCS data. The number of observations is 16,012 in all specifications. All specifications include a gender dummy and age-occupation-time(-gender) fixed effects. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C.5: Feminization of Occupations and the Effect on Tasks (2SLS, QCS Data – First Stage)

Dependent variable:	(1) Female Share	(2) Migrant Share
Female share instrument (f_{aot}^{IV})	0.512***	0.013
	(0.063)	(0.028)
Migrant share instrument (m_{aot}^{IV})	-0.210***	0.196***
	(0.043)	(0.036)
F-test:	44.57	15.03
SW F-test:	75.57	28.85

Notes: First-stage results for all five specifications shown in Table 4 in the main text. Pooled cross-sectional data of full-time workers from the QCS data. The number of observations is 16,012 in all specifications. As in Table 4, all specifications include a gender dummy and age-occupation-time(-gender) fixed effects. Standard errors (in parentheses) are clustered at the age-occupation-time level. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1.