Hidden in Plain Sight: Occupational Structure and the Gender Wage Gap

Riccarda Rosenball*[a], Tobias Eibinger†[a], Jörn Kleinert‡[a], and Ismir Mulalic§[b]

[a]Department of Economics, University of Graz ¶

[b]Department of Economics, Copenhagen Business School ‖

Abstract

We provide new evidence on how wage differences across firms shape the gender wage gap and highlight the role of gendered occupational segregation. We show that when occupations are omitted from canonical AKM models, the estimated firm effects mechanically absorb occupation-specific wage differences. We illustrate this mechanism through simulations and then quantify its empirical relevance using Danish register data. The estimated contribution of firm sorting to the gender wage gap declines by up to 30% once occupations are accounted for. Furthermore, around 90% of firm sorting occurs between industries, with women disproportionately employed in lower-paying ones. Our results demonstrate that gendered firm sorting is closely intertwined with occupational and industry segregation.

Keywords: wages, gender wage inequality, occupational segregation, AKM.
JEL codes: J16, J70, J81.

*riccarda.rosenball@uni-graz.at
†tobias.eibinger@uni-graz.at
‡joern.kleinert@uni-graz.at
§imu.eco@cbs.dk
¶Universitätsplatz 3, 8010 Graz, Austria.
‖Porcelænshaven 16A, DK-2000 Frederiksberg, Denmark.

1 Introduction

Earnings inequality between workers is largely explained by firm-specific differences in wage payments, often referred to as firm-specific wage premia (Abowd et al., 1999; Card et al., 2013; Song et al., 2018). These premia also constitute an important channel through which the gender wage gap is sustained. As conceptualized by Card et al. (2015), women are disproportionately employed in lower-paying firms – the "sorting effect"– and tend to receive lower wage premia than their male colleagues within the same firm, known as the "bargaining effect". Together, these two effects form a sizable firm component of the gender wage gap.[1]

In this paper, we identify gendered occupational segregation, whereby women and men are concentrated in occupations with different wage levels, as a central driver of gendered firm-sorting patterns. Because women and men work in different occupations, and therefore in firms with distinct occupational structures, the standard approach conflates firm and occupation effects, masking the role of gendered occupational segregation. We show, both conceptually and empirically, that once gendered occupational segregation is recognized as an independent channel contributing to the gender wage gap, part of what is commonly interpreted as firm-specific wage premia actually reflects differences in occupational structures across firms.

In Denmark, more than 80% of health care workers between 2013 and 2018 were women. The industry's average pay premium is below the overall average, and health care ranks third to last among the 36 industries in our data. Even within health care, women are concentrated in the lowest-paying jobs, such as care work, the lowest-paying 2-digit occupation among the 51 we observe. This pattern reflects sorting, but not necessarily into "low-paying firms". We focus instead on occupational segregation: women cluster in low-paying "female" occupations, while men sort into higher-paying, male-dominated occupations.

The standard framework to estimate firm-specific wage premia is the two-way fixed effects model introduced by Abowd, Kramarz and Margolis (1999) and Abowd et al. (2002), widely known as the AKM model. Wages are decomposed into additive worker fixed effects – capturing portable productivity premia – and firm fixed effects, reflecting persistent wage premia associated with specific firms and offered to all their workers.

The AKM framework has inspired extensive and influential research on wage setting. We extend this work by examining how occupations shape estimated firm effects. Our

[1] Quantitatively, the firm component accounts for approximately 15%–20% of the gender wage gap in Portugal and Italy (Card et al., 2015; Cardoso et al., 2016; Casarico and Lattanzio, 2024), 11 percent in France (Coudin et al., 2018), and as much as 40 percent in Estonia (Masso et al., 2022).

motivation is that the choices women and men make early in their careers – regarding education, skills, and occupational paths – are central to understanding gender differences in wages. Occupational choices determine the skills workers bring to firms, and these differ from general abilities: occupational skills are not portable across all firms to the same extent. A lawyer cannot do the work of a dentist. Moreover, firms differ in the types of occupations they demand. Occupations therefore form an important link between workers and firms, and recognizing this link is essential for interpreting firm effects in the AKM framework.

Estimating the AKM model is computationally demanding because it requires solving a massive high-dimensional fixed-effects problem. Adding occupation fixed effects not only requires detailed occupational data but also increases this computational complexity. Identification of firm fixed effects relies on worker mobility across firms. Some workers remain in their initial occupation all over their career, but many change both firm and occupation as circumstances evolve (Farber, 1999). Such changes are typically motivated by improved working conditions, career advancement, or higher pay. They may also reflect personal circumstances, such as the need for better work-life balance after childbirth (Manning and Petrongolo, 2008; Kleven et al., 2019; Card and Hyslop, 2021; Blundell et al., 2025; Bütikofer et al., 2025; Petrongolo and Ronchi, 2020; Borghorst et al., 2024). Many of these adjustments involve both a change of firm and a change in the occupation. In our data more than 30% of firm switchers also change the occupation.

This overlap between firm and occupation mobility makes it particularly important to control for occupations. Incorporating occupation fixed effects therefore allows to account for occupational heterogeneity more directly than approaches that estimate separate models for different occupational groups.

For example, Targa (2025) compare blue- and white-collar workers to assess whether these groups earn different firm pay premia within the same firm, and Card et al. (2015) estimate AKM models separately by occupational subgroups to examine whether the firm contribution to the gender wage gap varies between occupations. They find distinct patterns for different occupational groups. Our work is closer to another line of research that incorporates occupational heterogeneity more directly. Torres et al. (2018), Boza (2021), and Lhuillier (2025) include job-title fixed effects and show that occupational differences represent a distinct and substantial source of wage variation. Similarly, Cardoso et al. (2016) demonstrate that gender segregation across tasks accounts for as much of the gender wage gap as sorting across firms. Our contribution is to show that without accounting for occupations, firm fixed effects partly capture differences in occupational structures between firms. Once occupation controls are included, firm fixed effects measure wage premia net of a firm's occupational structure, so gendered firm sorting no longer reflects women's concentration in low-paying occupations or men's in high-paying ones.

To illustrate this mechanism, imagine two firms that are identical in their true wage-setting practices. Both firms pay according to the market rate for the occupations they employ and have a firm-level wage premium of zero. Firm A operates in care and administrative services. Most of its employees are women working in relatively low-paid occupations such as nursing assistants and office clerks. Because these occupations are generally paid less across the labor market, Firm A's average wage level is low. Firm B, by contrast, operates in engineering and software development. It employs mostly men in high-paying technical occupations such as mechanical designers and software engineers. Its overall average wage is therefore much higher, even though the firm itself does not pay a premium beyond what is typical for those occupations. When firm fixed effects are estimated without controlling for occupations, part of these occupational wage differences is attributed to the firms themselves. As a result, Firm A appears to have a low firm effect and Firm B a high one, even though their true firm-specific wage premia are identical. The apparent difference arises purely from their occupational structure rather than from differences in pay premia.

We show theoretically and through a simulation that occupational segregation causes the estimated firm contribution to the gender wage gap to partly reflect a firm's occupational structure. Without occupation controls, firm fixed effects capture an employment-weighted average of the omitted occupation effect. Because firms differ systematically in their occupational structure, those with a higher share of high-paying occupations tend to have higher estimated firm effects, while firms with more low-paying occupations, where women are overrepresented, appear to have lower ones. The simulation further demonstrates that the extent to which occupation effects are absorbed by firm fixed effects depends on the degree of occupational mobility: when workers frequently change occupations as they switch between firms, a larger share of the occupation effect is attributed to firms, whereas limited mobility shifts it toward the time-invariant worker fixed effects. When women are disproportionately employed in lower-paid occupations, firms with low estimated firm effects naturally exhibit higher female shares as well. Including occupation controls in the AKM model isolates firm wage premia net of occupational structure.

We use high-quality register data covering the universe of workers and firms in Denmark from 2008 to 2018. The dataset's detailed occupational codes enable us to examine how the gendered sorting component across firms changes when occupations are controlled for at varying levels of detail. We find that sorting into high-paying firms declines by about 20%-30% when occupations are taken into account. This suggests that standard AKM models partly attribute to firm sorting what is actually driven by women and men working in firms with different occupational structure. We further show that about 90%

of the remaining firm sorting component reflects women's disproportionate employment in lower-paying industries.

Beyond the AKM based research, the study refers to a range of literature. First, this paper contributes to the literature on the gender wage gap. While the gender wage gap has narrowed in recent decades due to convergence in education, skills, and labor market participation, women continue to earn significantly less than men (Altonji and Blank, 1999; Goldin, 2014; Goldin et al., 2017; Maasoumi and Wang, 2019). There is evidence that traditional human capital variables now account for only a small share of the gender wage gap, while occupational and industry segregation remain central drivers (Blau and Kahn, 2017), even for countries with low gender wage gaps like Denmark (Gallen et al., 2019). A large body of work has explored and identified several channels underlying this persistent gap, including the "child penalty", as childbirth induces career interruptions and reductions in labor supply (Kleven et al., 2019; Coudin et al., 2018; Kleven et al., 2024), as well as higher commuting costs for women, which restrict job search and result in shorter commutes (Manning and Petrongolo, 2008; Card and Hyslop, 2021; Le Barbanchon et al., 2021; Petrongolo and Ronchi, 2020; Borghorst et al., 2024). We show that occupational segregation also constitutes a systematic source of the gender wage gap.

Second, our paper contributes to the literature examining the role of firms in shaping wage inequality. Using the AKM framework, Card et al. (2013) show that wage-setting practices differ substantially across firms in Germany, and that high-wage firms (those with a high firm fixed effect) disproportionately employ high-wage workers (those with a high worker fixed effect), reinforcing overall wage inequality. Similarly, Song et al. (2018) find that in the U.S., firm-specific factors have been a central driver of the rise in earnings inequality since 1978, primarily through the increasing concentration of high-wage workers in high-wage firms. The sorting of high-wage workers into high-wage firms becomes even stronger once the limited-mobility bias is corrected using the approach proposed by Bonhomme et al. (2022).[2] Our contribution here is to offer the link through which high-wage workers are bound to high-wage firms: the occupation which is in high demand and therefore highly paid. Controlling for occupations removes the occupation-related component from both the worker fixed effects and the firm fixed effects.

Third, we contribute to the labour economics literature emphasizing the role of industry wage premia as an important source of wage inequality. Industry wage premia

[2] A well-known concern in this literature is the limited-mobility bias, which arises when firm fixed effects are estimated under sparse worker mobility between firms (Andrews et al., 2008, 2012). When most workers remain with the same firm or change jobs only within a narrow segment of the firm network, the estimated variance of firm effects tends to be downward biased, and the estimated correlation between worker and firm effects correspondingly attenuated.

vary substantially across industries and account for a significant share of overall wage inequality and the gender wage gap (Card et al., 2024; Buzaglo-Baris, 2025). Haltiwanger et al. (2022) show that recent increases in U.S. earnings inequality have been driven by a small set of industries at the top of the wage distribution – such as high-tech, STEM, finance, mining, and selected healthcare industries – and at the bottom including segments of retail and low-wage services. Since women are disproportionately concentrated in low-paying industries, this polarization may further deepen gender segregation in the labor market.[3,4]

The remainder of the paper is organized as follows. Section 2 outlines the theoretical motivation for incorporating occupational variables in the estimation of the AKM model. Section 3 presents the results from testing the theoretical predictions using Danish register data. We begin by describing the data in Section 3.1, followed by a discussion of key empirical facts in Section 3.2, and present and discuss the main empirical findings in Section 3.3. Finally, Section 4 concludes.

2 Occupational Segregation and the Gender Wage Gap

This section develops the theoretical motivation for incorporating occupations into AKM models. Section 2.1 introduces the AKM decomposition of the gender wage gap, and Section 2.2 explains how occupations shape both worker and firm fixed effects and, through this channel, contribute to gendered sorting across firms. Section 2.3 then presents a simulation study that shows that, when occupations are omitted, firm fixed effects absorb differences in firms' occupational structures, with the magnitude of this absorption depending on how often workers change occupations when switching between firms. Once gendered occupational segregation is added, this mechanism generates apparent patterns of gendered firm sorting, even when true firm premia do not differ by gender.

2.1 Decomposition of the Gender Wage Gap

The model developed by Abowd et al. (1999), commonly referred to as the AKM framework, uses linked firm-worker data to decompose wages into components attributable to both worker and firm characteristics. The AKM framework has become a widely used

[3] Palladino et al. (2025) finds that in France, the contribution of firms to the gender wage gap declined between 2002 and 2019, primarily due to a reduction in between-industry wage premia.

[4] Bruns (2019) documents that in Germany, the firm component of the gender wage gap increased markedly between 1990 and 2009, driven in part by declining union coverage and the expansion of firm-level collective bargaining.

tool in empirical labor economics, particularly for analyzing wage inequality and the role of firms in shaping gender wage gaps. In matrix form, the model can be expressed as

$$w = X\beta + D\alpha + F\psi + \varepsilon, \tag{1}$$

where w denotes the $N^* \times 1$ stacked vector of log wages across all worker-period observations, with N^* denoting the total number of worker-period observations. The matrix X is $N^* \times v$ and contains v time-varying worker characteristics (typically a full set of year dummies and flexible age controls interacted with education), with associated coefficients collected in the $v \times 1$ vector β. The matrices D and F are $N^* \times N$ and $N^* \times J$ indicator matrices for workers and firms, and α and ψ are the corresponding $N \times 1$ worker fixed effects (capturing time-invariant worker heterogeneity) and $J \times 1$ firm fixed effects (typically interpreted as firm-specific wage premiums); ε is an $N^* \times 1$ idiosyncratic error term.

The AKM framework relies on statistical assumptions ensuring that the error term is mean-independent and serially uncorrelated, after controlling for worker and firm IDs and observed worker characteristics. Economically, it assumes additive separability of worker and firm effects and exogenous worker mobility so that switches between firms are not driven by unobserved factors correlated with wages.[5] These conditions ensure that firm fixed effects capture persistent wage-setting differences rather than transitory shocks or selection effects.

An increasing number of studies follow Card et al. (2015) in estimating AKM models separately for women and men and use a simple mean decomposition in the second step to measure how gender differences in average firm fixed effects contribute to the gender wage gap. Specifically,

$$w \;=\; X\beta^G \;+\; D\alpha^G \;+\; F\psi^G \;+\; \varepsilon, \tag{2}$$

where each term is defined on the gender-specific sample G, and F is the firm indicator matrix from the full sample.

Card et al. (2015) argue convincingly that equation (2) "provides a simple framework for measuring the impact of firm-level pay premiums on the gender wage gap" (p. 649).

[5]The assumption of additivity in the Abowd–Kramarz–Margolis (AKM) model has been widely debated in the literature. Woodcock (2008) shows that omitting unobserved worker–firm match quality leads to an overstatement of both worker and firm contributions to wage inequality, as match effects are absorbed into the estimated fixed effects. Bonhomme et al. (2019) relax the additive separability assumption by allowing for interactions between worker and firm heterogeneity and find evidence of non-additive complementarities using Swedish register data. They find that the standard additive specification still approximates mean wage patterns reasonably well.

They decompose the difference in average firm pay premia between women and men into two components, *bargaining* and *sorting* effects, as follows:

$$\mathbb{E}\big[F_i\psi^M \mid \text{male}\big] \;-\; \mathbb{E}\big[F_i\psi^F \mid \text{female}\big] = \underbrace{\mathbb{E}\big[F_i(\psi^M - \psi^F) \mid \text{male}\big]}_{\text{bargaining effect}} + \underbrace{\mathbb{E}\big[F_i\psi^F \mid \text{male}\big] \;-\; \mathbb{E}\big[F_i\psi^F \mid \text{female}\big]}_{\text{sorting effect}}, \tag{3}$$

where the first term captures the "bargaining effect", i.e., the average difference in firm pay premia between women and men within firms, while the second term reflects the "sorting effect", i.e., differences arising from women and men working in firms with different average pay premia. The decomposition in equation (3) yields the sorting effect when firm fixed effects estimated from the female sample are used as the reference. We continue to refer to these as the "female firm fixed effects."[6]

2.2 Occupational Effects in Firm Fixed Effects

The AKM wage equation is typically estimated without accounting for systematic differences in wages across occupations. However, certain occupations consistently offer higher pay than others (Hsieh et al., 2019). We emphasize that occupation is a time-varying attribute of a worker that evolves over the course of a career. When workers change firms, they often also switch occupations, meaning that occupational assignment co-moves with firm assignment. To account for this, we extend the regression by introducing an $N^* \times o$ occupation-incidence matrix O, where o denotes the total number of occupations available. The extended wage equation becomes:

$$w \;=\; X\beta^G \;+\; D\alpha^G \;+\; F\psi^G \;+\; O\gamma^G \;+\; \varepsilon, \tag{4}$$

where γ^G is the vector of occupation-specific wage premia, estimated separately by gender.

Including occupation controls in the AKM regression allows us to distinguish occupation effects from worker and firm fixed effects, even within a fixed effects framework. Occupational changes can occur both when switching firms and when advancing within the same firm (e.g., promotion to management). Adding O directly affects the estimation of both worker fixed effects α^G and firm fixed effects ψ^G, as part of the variation previously attributed to these components is now captured by occupational differences. We

[6]Notice that the relative size of the bargaining and sorting components depends on the chosen reference. If "male firm effects" are the reference, then the decomposition is $\mathbb{E}\big[F_i\psi^M \mid \text{male}\big] - \mathbb{E}\big[F_i\psi^F \mid \text{female}\big] = \mathbb{E}\big[F_i(\psi^M - \psi^F) \mid \text{female}\big] + \mathbb{E}\big[F_i\psi^M \mid \text{male}\big] - \mathbb{E}\big[F_i\psi^M \mid \text{female}\big]$.

treat occupations as exogenous, that is assigned independently of firm characteristics, or unobserved wage determinants. Equation (4) describes the true wage-setting process as an AKM model that explicitly accounts for occupation effects. When occupations are omitted from the estimation, the estimated firm fixed effect captures both the pure firm effect and a weighted average of the occupation effects implied by the projection of $O\gamma^G$ onto the column space of F. The longer workers remain in a firm, and the larger the number of workers in specific occupations within that firm, the greater the influence of the uncontrolled occupation effects on the estimated firm fixed effect. The firm fixed effect can therefore be interpreted as a combination of the true firm premium and an employment-weighted average of the omitted occupation effects.

Formally, when estimating firm effects ψ^G without controlling for occupations within gender G, the estimator can be expressed as:

$$\hat{\psi}^G \; = \; \psi^G \; + \; (F'F)^{-1}F'O\gamma^G. \tag{5}$$

The second term on the right-hand side reflects the extent to which occupation effects are absorbed into the firm fixed effects. This depends on the correlation structure between firm assignments F and occupational structures O. If there is no systematic relationship between firms and occupations, meaning that each firm has a random mix of high- and low-paying jobs, the estimated firm effects will reflect the true firm-specific wage premia. However, in the presence of systematic occupational structures across firms, a substantial share of occupational wage differences will be mistakenly attributed to firms.

Similarly, the estimated worker fixed effects capture the time-invariant component of occupational effects when occupations are not controlled for in the estimation of the AKM equation:

$$\hat{\alpha} \; = \; \alpha \; + \; (D'D)^{-1}D'O\gamma^G. \tag{6}$$

The second term on the right-hand side equals zero only if workers are randomly assigned to occupations or if all occupations are paid the same.

We continue to conduct a Monte-Carlo simulation to assess how gendered occupational segregation, worker mobility, differences in occupational structures, and related factors contribute to the extent to which firm fixed effects reflect occupation-driven rather than firm-specific wage differences.

2.3 Simulation

The simulated data are generated from the general AKM model augmented with occupation effects and thus follow the standard AKM assumptions, while all additional assumptions in the simulation are kept as simple as possible. The data-generating process includes $N = 10,000$ workers and $J = 200$ firms observed over $T = 7$ periods. Each firm j employs an equal number of workers (N/J) but differs in its occupational structure. For simplicity, we define two occupations: low-pay $(o_{i(j,t)} = 0)$ and high-pay $(o_{i(j,t)} = 1)$. Firms vary in their share of high-pay positions, where each firm's high-pay share z_j is drawn from a uniform distribution $z_j \sim \mathcal{U}(0, 1)$. In total, there are as many low-paid as high-paid occupations needed in the whole economy. Firm capacities remain constant over time, ensuring that occupational structures are fixed across periods. The true firm fixed effects are set to zero for all firms $(\psi_j = 0)$ and years.

Half of the workforce is female and half is male. To maintain simplicity, the simulation does not include heterogeneity in observable worker characteristics such as education or experience, as these factors are typically controlled for in AKM applications. Each worker i, however, is assigned a true worker fixed effect α_i, drawn from a normal distribution with mean 0 and standard deviation 0.2, irrespective of gender. Women and men are assigned to occupations according to a gender segregation parameter s, which defines the ex-ante degree of gendered occupational segregation. Specifically, s measures the difference between the share of women in low-pay occupations and the share of men in low-pay occupations. The parameter ranges from 0, where women and men are equally likely to hold low-pay occupations, to 1, indicating full occupational segregation.[7,8]

In the initial period, workers are randomly assigned to firms, respecting each firm's fixed occupational structure. There is no unemployment and no unfilled vacancies, the total number of workers in low- and high-pay occupations exactly matches the available positions across firms.

In each subsequent period, the simulation follows a structured sequence. Following Andrews et al. (2008), each worker decides whether to change the firm based on a fixed probability, which we set to $p = 0.1$. For firm switchers, a parameter κ determines the probability of retaining their previous occupation.[9] After determining each switcher's firm and occupation preferences, a demand-supply matching process assigns workers to firms through a "job lottery" while firm's occupational structures are respected. If a worker intends to switch but no vacancies are available in the desired occupation, the worker remains in their current firm and occupation. This process is repeated across periods, generating a dynamic firm-worker panel with worker mobility and occupational

[7] The segregation parameter s is given by $s = \frac{W^{LP}}{W^{all}} - \frac{M^{LP}}{M^{all}}$, where W^{LP} denotes the number of women in the low-paid occupation and M^{LP} is the corresponding number for men.

[8] If the number of women is less than the number of low-pay positions, the remaining low-pay positions are filled by men, and no women occupy high-pay positions. If the number of women exceeds the number of low-pay positions, the remaining women are assigned to high-pay occupations.

[9] When $\kappa = 1$, all firm switchers keep their occupation; when $\kappa = 0$, all firm switchers change occupations. At $\kappa = 0.5$, half of the firm switchers change occupations when switching firms.

changes.

Wages in the simulation are generated according to a simple wage-setting rule. The level wage of worker i in period t, denoted by w_{it}, is given by:

$$w_{it} = 100 \cdot \gamma^{o_{i(j,t)}} \cdot \exp\left(\alpha_i + \varepsilon_{it}\right), \tag{7}$$

where α_i is the worker-specific effect, $o_{i(j,t)} \in \{0, 1\}$ indicates whether worker i, employed at firm j in period t, holds a high-pay occupation ($o_{i(j,t)} = 1$) or a low-pay occupation ($o_{i(j,t)} = 0$), and $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is a normally distributed idiosyncratic error term with standard deviation $\sigma_\varepsilon = 0.1$. The parameter γ denotes the occupation effect and is set to $\gamma = 1.5$, such that high-pay occupations pay 50% more than the base wage. The base wage for low-pay occupations is fixed at 100.

We then estimate a conventional AKM wage equation, omitting occupation controls. The estimated model is:

$$\ln w_{it} \;=\; \hat{\alpha}_i + \hat{\psi}_{j(i,t)}, \tag{8}$$

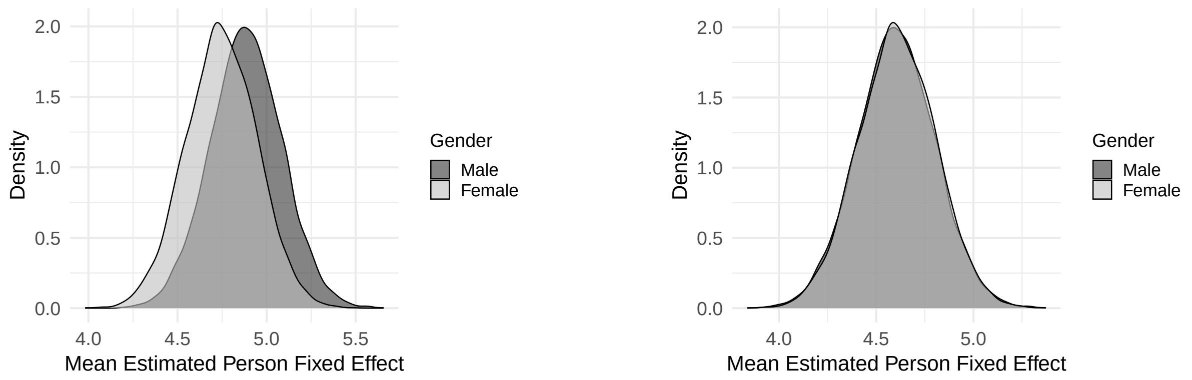
where $\hat{\alpha}_i$ denotes the estimated worker fixed effect, and $\hat{\psi}_{j(i,t)}$ is the estimated firm fixed effect for the firm j employing worker i in period t.

We perform 1,000 Monte Carlo replications, reassigning workers to firms and occupations and re-drawing wages in each iteration. We measure the extent to which occupational structure overstates firm fixed effects using the omitted variable bias formula, Deviation$(\hat{\psi}_j) = E[\hat{\psi}_j] - \psi_j$, with $\psi_j = 0$ by construction.

2.3.1 Occupation Effects in AKM Fixed Effects

Figure 1 shows the distribution of estimated worker fixed effects $\hat{\alpha}_i$ from the AKM model (8) without occupation controls (Figure 1a) and with occupation controls (Figure 1b). We first examine the case where the probability of retaining one's occupation when moving between firms is 50% ($\kappa = 0.5$), and gender segregation is apparent ($s = 0.5$), meaning that women are three times more likely than men to work in lower-paying occupations. Without occupation controls, the time-invariant component of the occupation effect is absorbed by the worker fixed effects. Because more women work in lower-paid occupations, their estimated worker effects are shifted to the left, while men's are shifted to the right. This creates gender differences in worker fixed effects, even though within-occupation wage setting is identical for women and men in our setup. In contrast, once occupation controls are included, the occupation component is removed from the worker fixed effects, and the distributions for women and men become identical.

Figure 1. Distribution of estimated worker fixed effects by gender and the relationship between estimated firm fixed effects and the female share of the firm.



(a) Without occupation controls

(b) With occupation controls

Note: The distribution of estimated worker fixed effects, $\hat{\alpha}_i$, by gender are given under the segregation parameter $s = 0.5$ and the probability $\kappa = 0.5$ to retain the occupation. Without occupation controls (panel a), occupational sorting shifts the distribution for females to the left. Apparent gender differences in worker effects are generated despite an identical within-occupation wage setting in the data-generating process. When occupational controls are included (panel b), the estimated distributions by gender coincide.

The time-varying component of the occupation effect, however, is absorbed by the estimated firm fixed effect. Firms with a higher share of high-paying occupations are attributed higher estimated firm fixed effects, even though no true firm-level wage differentials exist in the data-generating process. Table 1 shows that the share of high-paying occupations within a firm is strongly correlated with the estimated firm fixed effect (column (i)). Without occupation controls, the female share of the firm is also negatively related to the estimated firm fixed effect (column (ii)), reflecting women's disproportionate concentration in low-paying occupations under occupational segregation ($s = 0.5$). This negative relationship is entirely spurious in our simulation, as the female share merely proxies for the occupational structure of the firm. When occupations are controlled for in the AKM estimation (column (iv)), both spurious relationships disappear, indicating that the occupational effect is fully removed from the estimated firm fixed effects.

Table 1. Firm fixed effects estimates with and without occupation controls

	(i)	(ii)	(iii)	(iv)
	w/o occupation controls			w/ occupation
Intercept	−0.132***	0.291***	0.141***	−0.000
	(0.000)	(0.000)	(0.023)	(0.006)
High-pay share	0.252***		0.090***	0.000
	(0.000)		(0.014)	(0.004)
Female share		−0.582***	−0.375***	0.000
		(0.001)	(0.032)	(0.008)

Note: Estimates are from regressing the deviation in firm fixed effects on *High-Pay Share* and *Female Share* of the firm. Estimates are based on averages from 1,000 Monte Carlo replications. Standard errors in parentheses. All regressions use weighted least squares (WLS). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2 reports the simple mean decomposition of the gender wage gap into sorting and bargaining components for the benchmark simulation. The raw gap is about 0.1462 log points in our simulation, arising entirely from ex-ante gendered occupational segregation. The two rows show results from equation (3), once using firm fixed effects from a model without occupation controls and once with occupation controls.

When we estimate the AKM model without controlling for occupations separately for women and men and use equation (3) to decompose the difference in the gender-specific firm fixed effects, we find that firms appear to contribute 24% to the gender wage gap.[10] This shows that the omitted occupation effects are being absorbed into the firm fixed effects. The sorting component is positive, suggesting that women are incorrectly identified as clustering in lower-paying firms, even though the true driver is occupational segregation. The bargaining effect, as expected, is essentially zero since women and men receive identical firm-specific pay premia within firms in our simulation.

If we turn to the AKM with occupational controls case, which reflects the true structure of the simulation, both the sorting and bargaining effects disappear. In other words, once occupation is fully accounted for, firm fixed effects play no role in explaining the gender wage gap at all in our simulation.

2.3.2 The Role of Occupational Segregation and the Degree of Occupational Mobility

The extent to which we observe spurious gendered firm sorting in our simulation depends on the degree of gendered occupational segregation, captured by the segregation param-

[10]Note, to make firm fixed effects estimated separately for women and men comparable, we normalized the female firm effects by shifting their mean to match the male average.

Table 2. Decomposition of Firm Contribution to the Simulated Gender Wage Gap

Specification	Raw pay gap	Sorting	Bargaining	Total firm effect	% of pay gap
w/o occupation	0.1462	0.0348	0.0003	0.0351	24.0%
w/ occupation	0.1462	0.0000	0.0000	0.0000	0.0%

Note: The decomposition follows equation (3). Using firm fixed effects from an AKM model without occupational controls yields a spurious firm contribution to the gender wage gap, as omitted occupation effects load onto the firm component. When occupational controls are included, both sorting and bargaining components drop to zero, consistent with the structure of the simulation.

eter s. When there is no gendered occupational segregation ($s = 0$), women and men are equally represented across occupations and, consequently, not sorted across firms. Under complete segregation ($s = 1$), all women work in low-paying occupations and all men in high-paying ones, generating the maximum degree of spurious sorting: women are concentrated in firms with high female shares and low-paying occupations, while men are concentrated in firms with low female shares and high-paying occupations.

Figure 2 illustrates this pattern: the left column shows $s = 0$, the right column $s = 1$, with the top panel displaying the distribution of estimated worker effects by gender and the bottom panel plotting estimated firm fixed effects against female share within the firm.

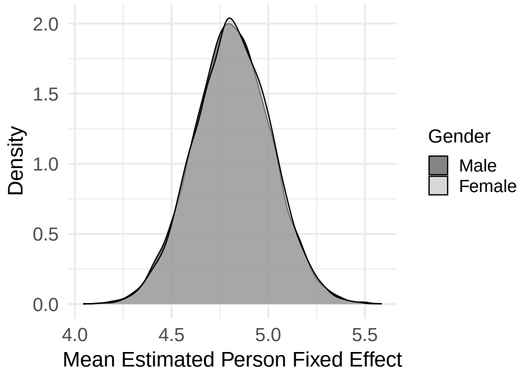
With no occupational segregation, the simulation yields no raw gender wage gap at all: worker fixed effect distributions are identical across genders, and firm effects show no relationship with the female share of the firm, since firms employ roughly equal shares of women and men. Under complete segregation, a large raw gender wage gap emerges, with pronounced gender differences in worker effects (Figure 2b) and a strong negative link between female share of the firm and the estimated firm fixed effects (Figure 2d), entirely driven by uncontrolled occupation effects.

While gender segregation (s) determines the size of the raw gender wage gap, the probability (κ) that a job switcher retains their occupation governs whether the occupation effect is absorbed more into the worker fixed effect or more into the firm fixed effect.

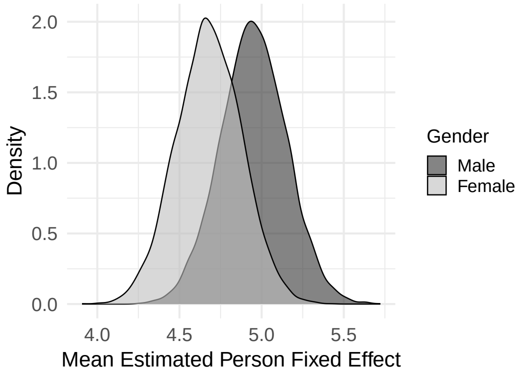
Figure 3 displays the average coefficients on the high-pay share and the female share of the firm from regressions predicting the estimated firm fixed effect, $\hat{\psi}_j$. Results are shown for $s = 0.5$ and varying degrees of occupational mobility, κ.

At lower values of κ, fewer workers change occupations when switching between firms, which shifts more of the occupation effect into the estimated firm effects. In this case, the estimated firm fixed effects become strongly related to firms' occupational structures: firms with a higher share of high-paying occupations display higher estimated firm effects. Because occupational segregation links occupation and gender, the female share of the

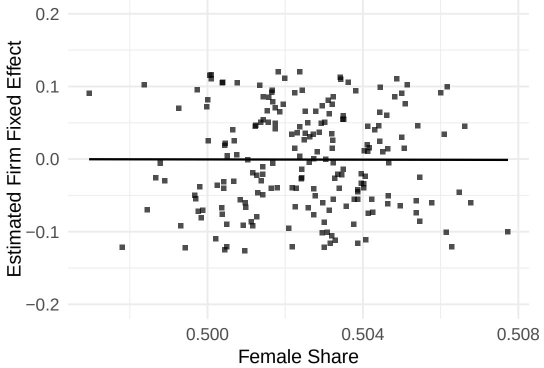
Figure 2. Distribution of mean estimated worker fixed effects by gender and firm fixed effects vs. female share of the firm.



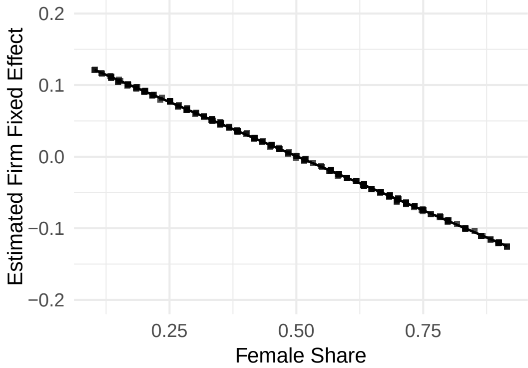
(a) No occupational segregation



(b) Full occupational segregation



(c) No occupational segregation



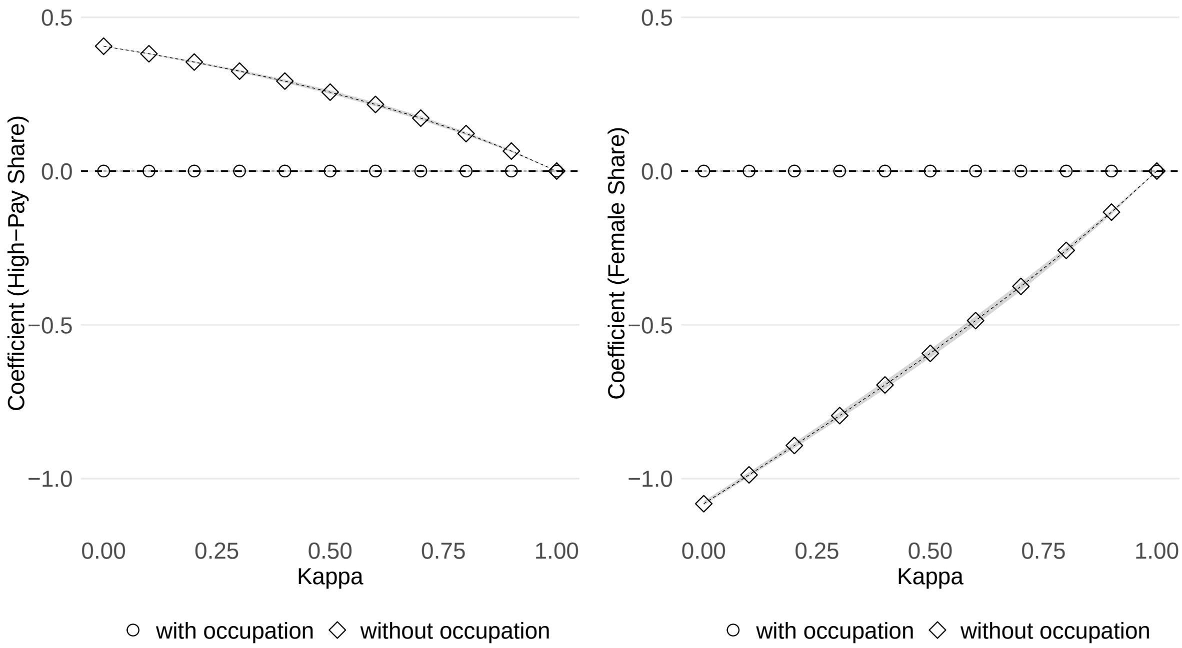
(d) Full occupational segregation

Note: Panels (a) and (c) correspond to no gendered occupational segregation ($s = 0$), and panels (b) and (d) to full segregation ($s = 1$). Occupational persistence is fixed at $\kappa = 0.5$, i.e., job switchers retain their occupation with probability 0.5. With $s = 0$, worker fixed effect distributions are identical across genders (panel a), and estimated firm effects show no relation to the female share (panel c). Under full segregation, pronounced gender differences in worker effects emerge (panel b), and estimated firm effects become strongly negatively associated with the female share (panel d), reflecting occupation-driven sorting rather than true firm heterogeneity.

firm becomes negatively correlated with the estimated firm effects, creating an apparent pattern in which female-dominated firms seem to pay less. This association, however, merely reflects occupational segregation rather than true firm-level pay differences.

Controlling for occupations in the AKM wage regression ensures that the estimated firm effects capture only the pure firm-specific pay component. This adjustment removes the spurious relationships between the estimated firm fixed effects and both the high-pay share and the female share of the firm. In both panels of Figure 3, the coefficients are close to zero for all values of occupational mobility κ, confirming that once occupation effects are controlled for, the spurious relationships disappear entirely.

Figure 3. Correlation between firms' occupational structure (high-pay share) and female share with the estimated firm fixed effect across different κ values.



(a) Correlation with high-pay share

(b) Correlation with female share

Note: Average coefficients from regressions of the estimated firm fixed effect on firms' occupational and gender compositions are reported across values of the occupational-persistence parameter κ. Panel (a) shows how low κ values generate strong spurious correlations between the estimated firm effects and the share of high-paying occupations, while panel (b) shows the corresponding spurious correlation with the female share of the firm. Once occupations are included in the AKM specification, both patterns disappear and the coefficients remain near zero for all κ. Results are based on 1000 bootstrap simulations.

3 Empirical Analysis

We now turn to the empirical analysis to bring the theoretical insights to the data. Our goal is to assess how strongly gendered occupational segregation contributes to the observed patterns of firm sorting. For this purpose, we use high-quality register data from Denmark.

3.1 Data

The empirical analysis draws on longitudinal administrative register data covering the entire working population in Denmark over the period 2008–2018. The data includes the universe of firms and their workers, providing detailed information on wages as well as an unusually rich set of covariates relative to standard register-based sources. For individual workers, the data contains information on age, gender, educational attainment, job tenure, industry of employment, and the occupation. Occupational classifications follow the Danish International Standard Classification of Occupations (DISCO), which allows us to group workers at varying levels of aggregation: 1-digit (10 groups), 2-digit (51 groups), 3-digit (171 groups), and 4-digit (550 groups).[11] The time-consistent personal identifier enables us to track individuals over time and to link information across multiple administrative sources, including tax records, and other relevant outcomes. For firms, we observe workforce size, and industry affiliation. Industries are classified up to the 3-digit level (127 industries) according to the Danish Industrial Classification (DB), which is directly aligned with the EU industry standard NACE.

We restrict the sample to all workers in Denmark aged 25–60 who are employed in firms with at least 10 workers. In addition, we impose standard sample selection criteria, excluding workers in the bottom and top percentiles of the wage distribution in order to remove extreme outliers. The final dataset comprises more than 12 million worker-year observations, covering 2.2 million distinct workers and 84 thousand firms. The dataset enables us to track job mobility, defined as movement from one firm to another.

Table 3 reports descriptive statistics for women and men in the periods 2008–2012 and 2013–2018. Women and men are similar in terms of mean age, accumulated experience, and number of children across both periods. Men, however, have a slightly higher probability of being highly educated. More importantly for our analysis, women earn approximately 20% lower wages than men. It is also noteworthy that, at the median, women tend to work in larger firms, and these firms have a substantially higher share of female co-workers.

Table 3. Summary statistics by gender and period (2008–2012 vs. 2013–2018)

	2008 - 2012		2013 - 2018	
	women	men	women	men
Age (years)	44	44	45	45
Experience (years)	22	22	23	23
Share highly educated (%)	15.9	18.6	20.4	21.5
Number kids	1.0	0.9	1.0	0.9
Hourly wage (DKK)	194.43	231.89	216.25	251.55
Firm size (median) (number of employees)	118	93	112	95
Fraction female co-workers (%)	35.7	20.1	36.3	19.9
Number firms	55,396	52,243	56,599	59,197
Number workers	804,171	902,635	860,943	964,275
Number of observations	2,604,767	2,993,116	3,183,804	3,679,918

Note: Wages are measured in DKK (1 DKK ≈ 0.15 USD).

[11]The Danish International Standard Classification of Occupations (DISCO) is the Danish version of the ISCO classification from the International Labour Organization.

3.2 Empirical Facts

From the simulation study above, we learned that the extent to which the estimated firm-sorting component declines when occupations are included depends on both the degree of occupational mobility and the extent of gendered occupational segregation.[12] To set the stage, we first document the frequency of occupational changes accompanying firm switches and the degree of gender segregation across occupations in the Danish labor market.

Occupational mobility Figure 4 illustrates how firm changes and occupational mobility are intertwined. The left panel traces, for each year between 2008 and 2018, the share of workers who switch firms (solid line) and the share of workers who simultaneously switch both firm and occupation at the 2-digit DISCO levels (dashed line). The patterns reveal that women and men are about equally likely to switch firms, yet women are less likely than men to change the occupation upon switching firms. Strikingly, occupational shifts accompany a large fraction of firm switches, typically more than one-third, and in some years even the majority. This underscores the close link between job mobility and occupational mobility in the Danish labor market.

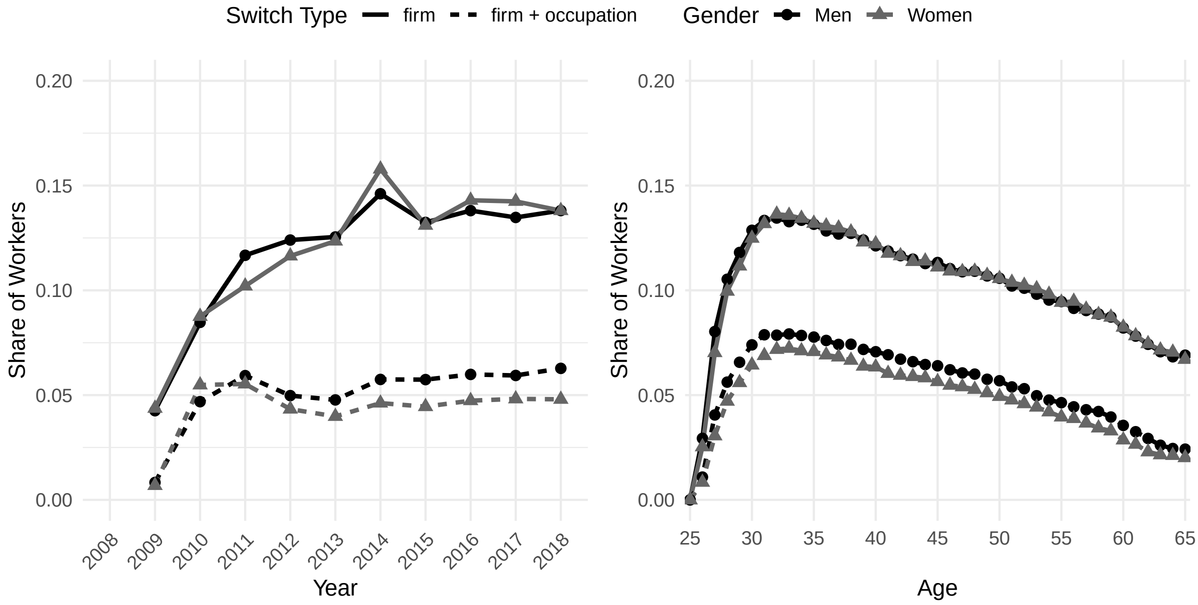
The right panel shows how firm and occupational mobility evolve over the life cycle, again separately for women and men. The likelihood of changing firms increases in the early career, peaking shortly after age 30, and then declines steadily with age. In this respect, women and men follow almost identical patterns. However, a different picture emerges when considering simultaneous changes of both the firm and the occupation. Here, the trajectories still peak after age 30 and taper off toward retirement, but a persistent gender gap opens up around age 28: women are consistently less likely than men to combine firm switches with occupational changes throughout their careers.

Occupational segregation To describe the extent of gender segregation across occupations in the Danish labor market, we draw on the Duncan index of dissimilarity (Duncan and Duncan, 1955).[13] Figure A.1 in the Appendix plots index values over time for different levels of occupational aggregation (DISCO 1–4 digits). At the 2-digit level, the index is slightly below 50%, while at the 3- and 4-digit levels it reaches around 50%.

[12] It also depends on firm heterogeneity in occupational structure and on occupational wage differentials. While we cannot summarize these dimensions in a single statistic at this stage, subsequent sections show that occupations differ markedly in their estimated occupation fixed effects and that firms employ very different mixes of occupations.

[13] The Duncan index of dissimilarity is a statistical measure of segregation that shows how evenly two groups are distributed across different areas. It ranges from 0 (complete integration) to 1 (complete segregation) and indicates the proportion of one group that would need to move for the groups to be evenly distributed.

Figure 4. Worker mobility over time and by age



Note: The left panel presents annual probabilities of changing firms and of changing both firm and (2-digit) occupation for women and men from 2008 to 2018. The right panel shows the corresponding probabilities by age.

These values point to a high degree of occupational segregation: depending on the level of detail in occupational coding, between 40% and 50% of either men or women would need to change occupations in order to achieve an equal gender distribution across occupations.

There is substantial variation in female representation, both across and within broad occupational groups. Figure A.2 in the Appendix illustrates the share of women across occupations at the one- and two-digit DISCO levels. This variation underscores the importance of using sufficiently detailed occupational codes in the analysis, as broader categories obscure meaningful differences in gender composition.

At the 1-digit level, female representation is particularly low in military occupations (0), craft and related trades (7), plant and machine operators and assemblers (8), and skilled agricultural work (6). By contrast, women are more strongly represented in service and sales (5) as well as in clerical support occupations (4). Managerial positions (1) remain predominantly male-dominated, with female shares often below 30%.

Looking more closely at the 2-digit level, women are highly concentrated in care work (53) and health occupations (22), in manual food preparation (94), cleaning (91), and in clerical support roles such as general office, secretarial, or customer service work, where female shares frequently exceed 80%. Some professional occupations also exhibit high female representation, including teaching (23), health professionals (22), and legal, social, and cultural professionals (26), with female shares above 80%. Among associate professionals, the female share likewise exceeds 70% in several groups, such as health technicians and assistants (32), business services, economics, administration, and sales (33), and legal, social, and cultural associate professionals (34).

3.3 Empirical Results: Occupational Segregation and the Gender Wage Gap in Denmark

We next use the Danish data to assess how controlling for occupation affects the estimated firm fixed effects and the extent of gendered firm sorting. We use the mean decomposition in equation (3) to see by how much the firm-sorting component declines when AKM firm fixed effects are estimated with occupation controls.

We estimate the AKM model once without occupation controls and then four times including occupational controls at the 1 to 4-digit DISCO levels, using log hourly wages separately for women and men in two periods (2008–2012 and 2013–2018). In addition to the occupational controls, we include the standard control variables: a full set of year dummies and flexible age controls interacted with education.[14] Following the literature, we normalize firm fixed effects to the hotel and restaurant industry to make gender-specific firm effects comparable. This normalization affects only the bargaining component, while the sorting component is invariant to the choice of the reference industry.

The density plots in Figure B.1 in Appendix B show the distribution of firm fixed effects estimated with and without occupation controls at the 2-digit and 4-digit DISCO levels for women and men during the 2013–2018 period. Overall, the two curves align closely across panels. Yet, the distributions with occupation controls are slightly more concentrated around zero. This pattern suggests that some of the dispersion in firm fixed effects without occupation controls reflects differences in firms' occupational structure. The adjusted estimates therefore capture firm premia net of occupational structure. While not crucial in many AKM settings, considering occupations is still informative, especially for understanding gendered firm-sorting patterns.[15]

[14] We estimate the model using the high-dimensional fixed effects algorithm proposed by Guimarães and Portugal (2011). Prior to estimation, we construct the largest connected sets of workers and firms separately for women and men. We then estimate the AKM models within each gender-specific connected set and restrict the comparison of firm effects to the intersection of the two sets.

[15] Worker fixed effects also capture the time-invariant component of the occupation in which a worker is employed. As shown in Figure B.2, once occupational effects are controlled for in the AKM model, the distribution of worker fixed effects becomes slightly more compressed around zero compared to the firm fixed effects from AKM without occupation controls.

3.3.1 Decomposition

Table 4 reports the main result of our paper: the decomposition of the gender wage gap into sorting and bargaining components using different levels of occupational controls. For the decomposition we are using female firm fixed effects as the reference. In the later period, 2013–2018 (Panel B), the overall gender wage gap is 0.1402 log points, with 8.3% attributable to sorting across firms. Accounting for occupations in AKM reduces this share step by step with the level of detail in the occupational codes, falling to 7.8% with 1-digit controls and to 5.8% with 4-digit controls. Including occupation controls reduces the estimated firm-sorting component by almost one third, indicating that occupational structure contributes meaningfully to what is otherwise attributed to firm sorting. A similar pattern is evident in the earlier period, 2008–2012, when the raw gap was larger with 0.1683 log points. The sorting share declines from 11.0% in the baseline AKM specification to 9.1% with 4-digit occupational controls, a reduction of nearly one fifth.[16]

These results show that controlling for detailed occupations reduces the estimated contribution of firm sorting to the gender wage gap by up to 30% in the Danish data. Importantly, this decline is robust to whether female or male firm fixed effects are used as the reference (see Appendix Table B.1).

The sorting effect declines once occupations are controlled for, indicating that part of what standard AKM models attribute to firm sorting is driven by women and men working in firms with different occupational structure. The reduction in the sorting component is around two percentage points, from 7.5% to 5.1 percent in 2013-2018 and from 10.0% to 8.3% in 2008–2012. The decline appears systematically across both time periods and is stronger with increasing occupational granularity.

Women change occupations less frequently than men and are less likely to move into higher-ranked occupations. In our data, only 6.6% of women switch occupations when changing firms, compared with 8% of men. Among these movers, women transition upward—into occupations with higher occupational fixed effects—at a probability of 3.3%, which is lower than the 4.2% observed for men. Their likelihood of moving downward is similarly lower (3.2% vs. 3.8% for men). The mobility patterns of women differ markedly from those of men.

These distinct gendered patterns in occupational mobility are illustrated in Figure 5,

[16] While our primary focus is on the sorting effect, we also observe adjustments in the bargaining effect. First, because women's occupational mobility is slightly lower than men's, less of the occupation effect is absorbed by the female firm fixed effects. Second, when women within a firm are concentrated in lower-paying roles, the firm's occupational structure appears to have a higher share of low-paying occupations in the AKM estimation for the female sample. As a result, the estimated firm fixed effects for a similar firm are higher for men and lower for women in the gender-specific AKM estimation without occupation controls.

Table 4. Decomposition of the gender wage gap (GWG) by period and specification

Specification	DISCO	Raw GWG	Sorting	Bargaining	Firm	Firm/GWG	Sorting/GWG
Panel A: 2008–2012							
W/o occupation	–	0.1683	0.0185	0.0061	0.0245	14.6%	11.0%
W/ occupation	1-digit	0.1683	0.0179	0.0060	0.0239	14.2%	10.6%
W/ occupation	2-digit	0.1683	0.0172	0.0057	0.0229	13.6%	10.2%
W/ occupation	3-digit	0.1683	0.0166	0.0047	0.0212	12.6%	9.8%
W/ occupation	4-digit	0.1683	0.0154	0.0084	0.0238	14.2%	9.1%
Panel B: 2013–2018							
W/o occupation	–	0.1402	0.0116	0.0009	0.0125	8.9%	8.3%
W/ occupation	1-digit	0.1402	0.0109	0.0010	0.0119	8.5%	7.8%
W/ occupation	2-digit	0.1402	0.0092	−0.0028	0.0065	4.6%	6.6%
W/ occupation	3-digit	0.1402	0.0091	−0.0005	0.0086	6.1%	6.5%
W/ occupation	4-digit	0.1402	0.0082	0.0015	0.0096	6.9%	5.8%

Note: Decomposition of the gender wage gap into sorting and bargaining components using female firm fixed effects as the reference group. Estimates are reported for two periods (2008–2012 and 2013–2018) and for AKM specifications with increasing levels of occupational detail. "Firm" denotes the firm-specific component of the decomposition, and the percentages express each component relative to the raw gender wage gap.

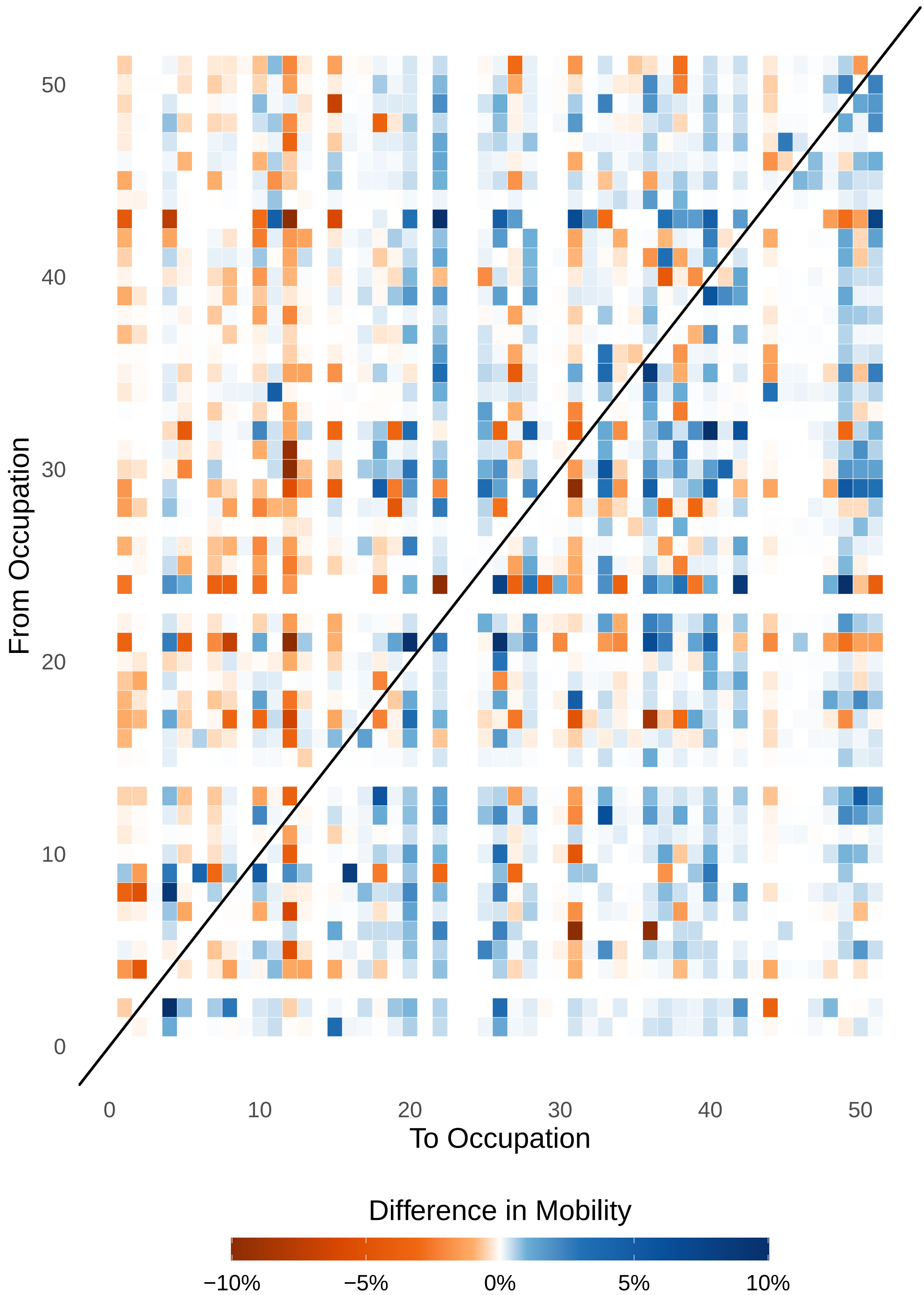
which shows the gender differences in the direction of occupational mobility. It illustrates transitions across 2-digit DISCO occupations between 2013 and 2018 and highlights how these patterns differ for women and men. Each cell reports the difference in transition probabilities between men and women when switching from the occupation on the vertical axis to the occupation on the horizontal axis. The shading reflects the difference in transition probabilities between men and women for each origin–destination pair, where orange cells indicate transitions that women make more frequently than men, and blue cells indicate transitions that men are more likely to make. Occupations are ordered along both axes according to their estimated occupational fixed effects coming from the AKM wage equation for men. Low-paying occupations appear on the left and bottom, while high-paying occupations appear on the right and top.[17,18] Importantly, transitions located below the main diagonal correspond to upward moves in terms of occupational fixed effects, while those above the diagonal represent moves into lower-paying occupations.

Figure 5 shows that women's occupational mobility is strongly directed toward the lowest-ranked occupations, with substantial inflows coming from a broad range of origin occupations spanning the entire distribution of occupational fixed effects. Women frequently transition into care work (rank 1), manual food preparation (rank 2), general office and customer service work (rank 5 and rank 7), cleaning work (rank 8) or general calculation and registration work (rank 10). Women exhibit strong mobility into clerical

[17] For the complete mapping of occupations to ranks, see Table B.3 in Appendix B.

[18] The diagonal corresponds to workers remaining in the same occupation and is omitted from the color scale to highlight occupational transitions.

Figure 5. Gender Differences in transition probabilities across occupations



Note: Each cell shows the difference in transition probabilities between women and men when switching from the 2-digit occupation on the vertical axis to the 2-digit occupation on the horizontal axis. Orange cells indicate transitions more frequent among women, while blue cells indicate transitions more frequent among men. Transitions are computed using worker mobility observed between 2013 and 2018.

and service-related occupations, such as secretarial and general office work (rank 12) and teaching and educational work (rank 15). These inflows originate from a wide range of initial occupations. In the mid-ranked part of the distribution, women are more likely than men to move into business services, economics, administration, and sales (rank 31), assistant work in health (rank 31), and, to some extent, into higher-ranked professional jobs in economics, administration and sales (rank 38) and work in health sector (rank 44). By contrast, the blue areas near the top of the ranking indicate that men are more likely than women to move into high-ranked managerial occupations (ranks 48–51). While some women do transition into management, these moves are less frequent than for men and, when they occur, they tend to be directed toward hotel and restaurant management or administrative management rather than senior management roles.

For men, the blue cells below the main diagonal indicate that they are relatively more likely than women to move into higher-ranked occupations. These upward differences appear most strongly for transitions originating from occupations ranked roughly 10–30 and leading into occupations ranked 20–40. Men more frequently move into mid- and upper-mid–ranked occupations such as engineering and natural sciences (rank 36), craft-based work in construction (rank 42), operator work in machinery (rank 37), assembly work (rank 39) and metal and machinery work (rank 40). The strongest gender differences occur at the top of the occupational ranking (ranks 47–51), where men are far more likely than women to move into management and military officer positions, drawing from a wide range of origin occupations.

3.3.2 Industry Pay Premium

Industries differ markedly in their occupational structure. As a result, women and men are segregated not only across occupations, but also sorted across industries. Figure A.3 in the Appendix illustrates the pronounced segregation across industries in Denmark, where the highest female shares are found in social institutions, followed by healthcare, education, other services, pharmaceuticals, public administration, and the hotel and restaurant industry. Prior work has emphasized the importance of such segregation patterns in explaining gender wage gap (Blau and Kahn, 2017). Petrongolo and Ronchi (2020) convincingly argue that occupational and industry segregation remain central drivers of the gender wage gap. Building on this insight, we construct industry wage premia and examine the extent to which between-industry sorting of women and men is captured by the firm sorting component. To construct the industry wage premia from firm fixed effects from AKM equations with occupation controls on the 4-digit level.

We follow Card et al. (2024) in measuring industry wage premia as the employment-weighted average of firm fixed effects i.e.

$$\Sigma_s^F = \frac{\sum_{j \in s} n_j \, \psi_j^F}{\sum_{j \in s} n_j},$$

where n_j denotes the number of person-year observations in firm j, and ψ_j^F the female firm effect of firm j.

To analyze the role of gendered sorting across industries in the gender wage gap, we decompose gender differences in firm fixed effects into a between- and a within-industry component. Let Σ^F denote the vector of industry pay premia, and let S be the firm-to-industry mapping matrix. With this notation in place, the between-firm component of the gender wage gap (sorting) can be formally expressed as

$$\underbrace{E[F_i\psi^F \mid \text{male}] - E[F_i\psi^F \mid \text{female}]}_{\text{between-firm gap}} = \underbrace{E[S_i\Sigma^F \mid \text{m}] - E[S_i\Sigma^F \mid \text{f}]}_{\text{between-industry}} + \underbrace{E[F_i\psi^F - S_i\Sigma^F \mid \text{male}] - E[F_i\psi^F - S_i\Sigma^F \mid \text{female}]}_{\text{within-industry}}.$$

The first term on the right-hand side captures differences in the industry distribution of women and men (between-industry sorting), while the second term reflects differences in the allocation of women and men across firms within the same industry (within-industry sorting).

Table 5 shows a decomposition of the between-firm component of the gender wage gap (sorting) into between- and within-industry contributions, based on a detailed classification of 128 industries available in our sample and using the female firm fixed effect as reference. The results show that roughly 90% of the sorting component is accounted for by between-industry sorting. This share is remarkably stable whether we use industry pay premia derived from AKM firm fixed effects with or without occupation controls. However, because controlling for occupations reduces the overall size of the firm-sorting component, the absolute magnitude of the between-industry contribution also becomes smaller. In other words, most of the gender differences in firm wage premia attributed to sorting can be traced back to the unequal distribution of women and men across industries. The remaining 10% reflect within-industry sorting, meaning that women and men are also allocated differently across firms operating in the same industry. Note, the results are consistent when using male firm fixed effects as the reference (see Table B.2 in the Appendix).

Similar to occupational mobility, we also see pronounced gender differences in industry mobility patterns. Figure 6 illustrates gender differences in patterns of mobility across

Table 5. Decomposition of firm sorting into between- and within-industry components (female firm fixed effects reference)

Specification	DISCO	Raw GWG	Sorting	Between ind.	Within ind.	Share btw.	Share within
Panel A: 2008–2012							
w/o occupation	-	0.1683	0.0185	0.0165	0.0020	89.3%	10.7%
w/ occupation	1-digit	0.1683	0.0179	0.0159	0.0020	88.9%	11.1%
w/ occupation	2-digit	0.1683	0.0172	0.0152	0.0020	88.6%	11.4%
w/ occupation	3-digit	0.1683	0.0166	0.0147	0.0019	88.5%	11.5%
w/ occupation	4-digit	0.1683	0.0154	0.0135	0.0019	87.8%	12.2%
Panel B: 2013–2018							
w/o occupation	-	0.1402	0.0116	0.0105	0.0011	90.3%	9.7%
w/ occupation	1-digit	0.1402	0.0109	0.0099	0.0010	90.6%	9.4%
w/ occupation	2-digit	0.1402	0.0092	0.0083	0.0010	89.6%	10.4%
w/ occupation	3-digit	0.1402	0.0091	0.0081	0.0009	89.5%	10.5%
w/ occupation	4-digit	0.1402	0.0082	0.0072	0.0010	87.8%	12.2%

Note: Decomposition of the firm sorting component of the gender wage gap into between- and within-industry sorting based on 128 industry categories. The between-industry term captures gender differences in the distribution of workers across industries, while the within-industry term reflects gender sorting across firms operating within the same industry. Estimates use female firm fixed effects as the reference. The percentages express the between-industry (btw.) and within-industry (within) component as a share of the sorting component.

36 sub-industries, which are ranked by their industry pay premia.[19,20]

The heatmap indicates that women's industry moves follow clear patterns. The orange vertical stripes show that many of their transitions lead to the same destination industries. Women's mobility is particularly directed toward lower-ranked industries such as public administration (PA), healthcare (HC), and education (E), but also toward mid-ranked industries including wholesale and retail trade (WR) and social institutions (SI). At the upper end of the ranking, women frequently move into legal and accounting consultancy (LAC), financial and insurance activities (FI) and the manufacture of pharmaceuticals (PH).

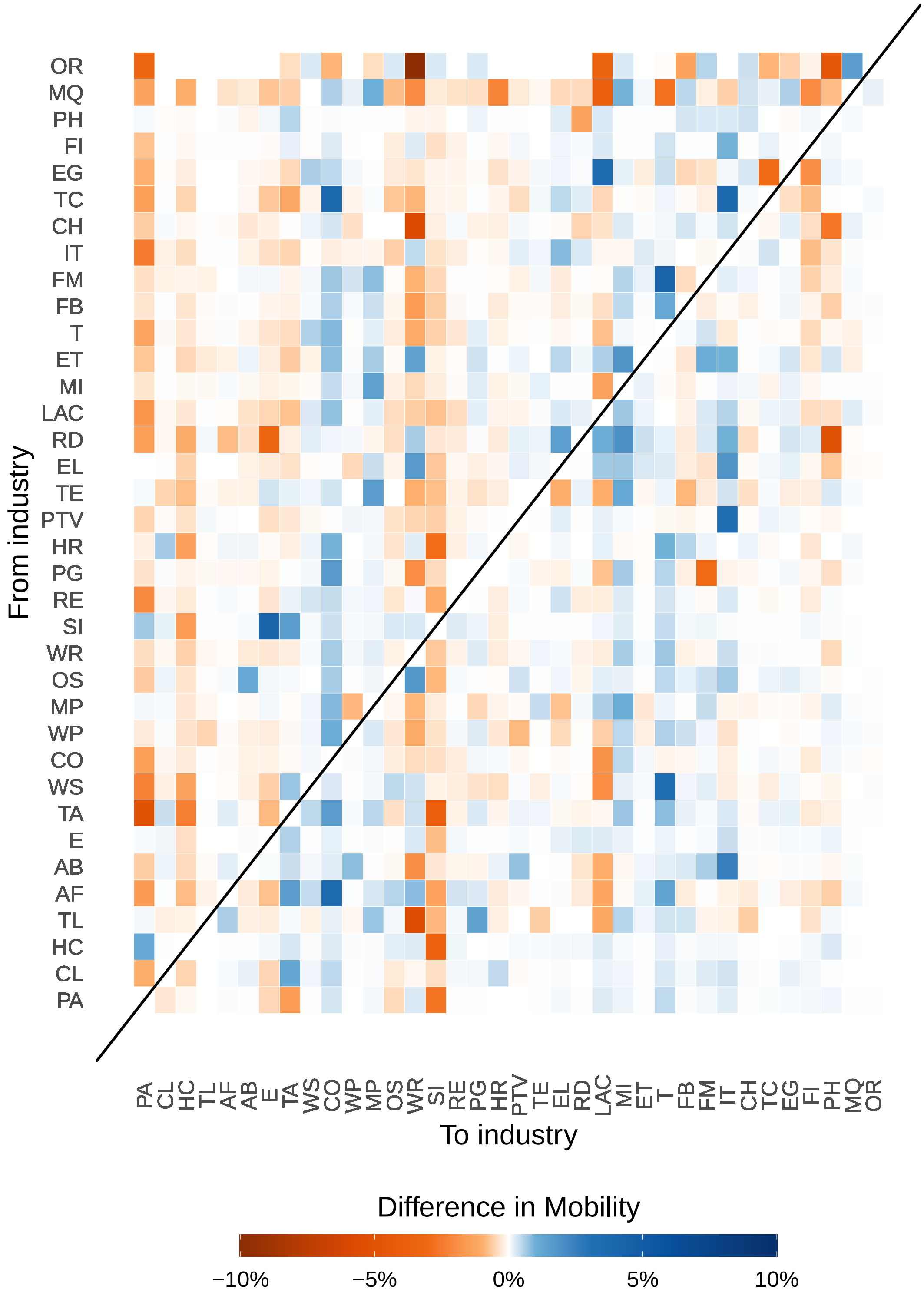
Men's mobility patterns are directed toward traditionally male-dominated industries, with strong transitions into metal production (MP) and construction (CO) in the middle-left of the ranking, as well as into mid- to high-ranked industries such as the machinery industry (MI), transport (T), and information technology (IT).

The patterns discussed above already reveal substantial gender differences in mobility

[19]The full mapping of industries to short codes, pay premia, and ranks provided in Table B.4 in Appendix B.

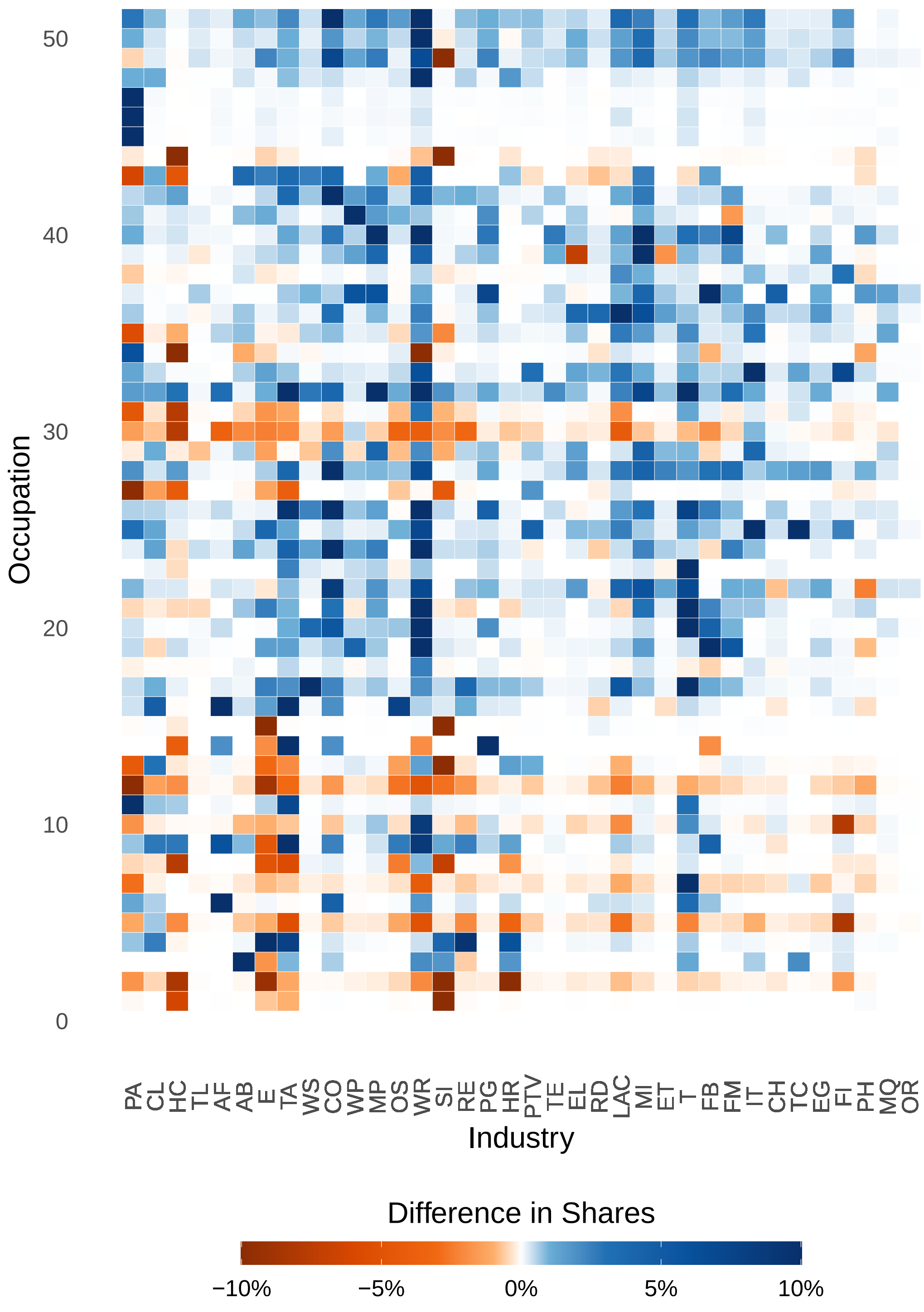
[20]Similar to occupational mobility from before, women are also less likely than men to switch to another industry (4.6% vs. 5.1%). For both genders, a slight majority of these moves are into lower-ranked industries (about 51%), but we see clear gender differences in the destination industry.

Figure 6. Gender differences in transition probabilities across industries



Note: Each cell shows the difference in transition probabilities between women and men when switching from the industry on the vertical axis to the industry on the horizontal axis. Orange cells indicate transitions more frequent among women, while blue cells indicate transitions more frequent among men. Transitions are computed using worker mobility observed between the 36 sub-industries from 2013 to 2018. Labels for industry abbreviations are provided in Table B.4 in Appendix B.

Figure 7. Gender Differences in the Distribution of Workers Across Industries by Occupation



Note: Each cell shows the difference in the share of workers (men–women) employed in a given occupation–industry combination. Positive values (blue) indicate industries where men form a larger share within that occupation; negative values (orange) indicate those where women are more represented. White cells arise when the cell has very few workers or when the male–female difference is close to zero. Labels for the 36 industry abbreviations are provided in Table B.4 in Appendix B.

across occupations and industry. Equally important is where workers end up in the occupational–industry landscape. To understand the actual allocation of workers across the labor market, we next examine how occupations and industries combine to form the jobs that women and men hold.

Women remain concentrated in both lower-paying occupations and lower-paying industries. From Figure B.3 in the Appendix, we observe a pronounced negative correlation of about -0.3 between the female share across 2-digit occupations and the corresponding occupation fixed effects, as well as between the female share across 36 sub-industries and the industry pay premia. This indicates that women are increasingly represented in lower-paying occupations and industries, a pattern that we can study in more detail in Figure 7. This heatmap shows, for each occupation, the difference between women and men in the share of workers allocated to each industry. Positive values (men–women) indicate occupation–industry combinations where men are more concentrated, while negative values indicate those where women are more concentrated. Occupations and industries are ordered according to their occupational fixed effects and industry pay premia.

The figure shows that women are overrepresented in the lowest-paying occupations (ranks 0–15). Several occupations—such as manual food preparation (2), customer services (5), other general office and customer service work (7), and general office and secretarial work (12)—are relatively low paid and female-dominated across nearly all industries. Women also dominate clerical support work (30) and occupations in business services, economics, administration, and sales, although these tend to rank higher in the occupational distribution. Men, by contrast, dominate higher-paying occupations and are more likely to work in high-paying industries, as well as in the better-paid occupations within lower-paying industries. This pattern is especially pronounced at the top of the occupational ranking, where management positions are overwhelmingly held by men across almost all industries.

Taken together, we reveal a consistent pattern of gender segmentation in the Danish labor market. Women's mobility is disproportionately directed toward lower-paying occupations and industries, while men are more likely to move into higher-ranked positions. These mobility patterns translate into a highly uneven allocation of women and men across occupation–industry combinations. Women cluster in low-wage occupations across a broad set of industries, whereas men dominate the higher-paying occupations and the more lucrative industries, and tend to hold the better-paid jobs even within lower-paying sectors. These patterns complement our decomposition results, which suggest that firm sorting is closely intertwined with occupational and industry segregation.

4 Conclusion

This paper examines occupational segregation – whereby women are disproportionately employed in lower-paying occupations and men in higher-paying ones – as a mechanism behind the empirical finding that women more often work in lower-paying firms. To address this, we revisit the widely applied AKM framework and look at changes in the firm sorting component when occupations are included as control variables in the AKM equation. Our results show that controlling for detailed occupations isolates firm wage premia net of occupational structure and significantly reduces the estimated contribution of firm sorting to the gender wage gap. Our analysis further suggests that detailed occupational controls matter: one-digit codes are too coarse, whereas two-digit classifications already capture much of the relevant variation.

We show theoretically and in a simple simulation that firms with a higher share of high-paying occupations appear to have higher estimated firm effects, while firms with more low-paying occupations tend to have lower ones. The extent to which occupation effects are absorbed into firm fixed effects depends on the degree of occupational mobility. When workers frequently change occupations upon switching between firms, the occupation effect is increasingly attributed to firms. When women are additionally concentrated in lower-paying occupations, this mechanism causes firms with high female shares to appear as low-paying, thereby inflating the estimated contribution of firm sorting to the gender wage gap. Considering this mechanism is important, as occupations differ markedly in pay, as reflected in their occupation fixed effects, and firms differ substantially in the occupational structures they employ across industries.

We then use high-quality Danish register data to assess the empirical relevance of this mechanism. In Denmark, too, we observe pronounced occupational segregation as well as frequent occupational changes when workers switch firms. We find that the estimated contribution of firm sorting declines up to 30% when occupational differences are taken into account in the Danish register data. These findings may be even more pronounced in labor markets with greater occupational wage differences and stronger gender segregation between occupations.

We further find that about 90% of firm sorting can be attributed to between-industry sorting of women and men. Since industries pay different wage premia, and women are disproportionately employed in lower-paying industries – mirroring occupational segregation – industry and occupational segregation emerge as central mechanisms underlying gendered firm sorting.

Gendered mobility patterns reinforce gender segregation in the labor market: women tend to move to lower-paying occupations and to industries offering lower wage premia, whereas men more frequently transition upward across both occupations and industries. We also document a combined effect: women are more likely to cluster in low-paying occupations across a broad range of industries, while men dominate higher-paying occupations and high-paying industries, and even within lower-paying industries, men tend to occupy the better-paid positions. These patterns confirm that gendered firm sorting is closely tied to occupational and industry segregation.

Our results reveal patterns of segregation in the Danish labor market, and several mechanisms highlighted in prior research offer plausible explanations for why such patterns emerge. Women may face barriers to accessing high-paying firms, limiting their opportunities to move into the upper end of the wage distribution (Sorkin, 2017; Casarico and Lattanzio, 2024). Workplace environments can also play a role. Harassment risks are higher in male-dominated and high-paying firms, which may discourage women from entering or remaining in these settings (Folke and Rickne, 2022; Chikhale et al., 2025). Beyond workplace conditions, preferences and constraints outside the labor market matter as well. Women often place greater value on flexibility, predictable hours, and other non-wage amenities due to different time constraints and caregiving responsibilities (Sorkin, 2018; Morchio and Moser, 2024; Labanca and Pozzoli, 2022). They are also more likely to seek out family-friendly jobs (Hotz et al., 2018) or positions closer to home that reduce commuting burdens (Bütikofer et al., 2025; Petrongolo and Ronchi, 2020; Borghorst et al., 2024).

These insights are relevant for policy debates on gender wage inequality and suggest that incorporating occupations can complement analyses focused only on within-firm or within-industry pay differences.

Acknowledgments The authors would like to thank Jos van Ommeren for useful suggestions on earlier drafts. They are grateful to Statistics Denmark for providing the data. The usual disclaimer applies.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 3 1999, *67* (2), 251–333.

_ **, Robert H. Creecy, and Francis Kramarz**, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Technical Paper TP-2002-06, Center for Economic Studies, U.S. Census Bureau 2002.

Altonji, Joseph G. and Rebecca M. Blank, "Chapter 48 Race and gender in the labor market," in "Handbook of Labor Economics," Vol. 3, Elsevier, 1999, pp. 3143–3259.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Andrews, M.J., L. Gill, T. Schank, and R. Upward, "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Economics Letters*, December 2012, *117* (3), 824–827.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet, "Gender differences in job search: Trading off commute against wage," *The Quarterly Journal of Economics*, 2021, *136* (1), 381–426.

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

Blundell, Richard, Hugo Lopez, and James P. Ziliak, "Labor Market Inequality and the Changing Life Cycle Profile of Male and Female Wages," *American Economic Journal: Applied Economics*, October 2025, *17* (4), 100–133.

Bonhomme, Stephane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler, "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?," *Journal of Labor Economics*, mar 2022.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Borghorst, Malte, Ismir Mulalic, and Jos van Ommeren, "Commuting, gender and children," *Journal of Urban Economics*, 2024, *144*, 103709.

Boza, István, "Wage Structure and Inequality: The Role of Observed and Unobserved Heterogeneity," CERS-IE Working Paper 2131, Institute of Economics, Centre for Economic and Regional Studies 2021.

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

Buzaglo-Baris, Shira, "Firms, industries and the gender wage gap," *Labour Economics*, December 2025, *97*, 102817.

Bütikofer, Aline, René Karadakic, and Alexander Willén, "Parenthood and the gender gap in commuting," *Journal of Public Economics*, August 2025, *248*, 105371.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *The Quarterly Journal of Economics*, 10 2015, *131* (2), 633–686.

_ and Dean R. Hyslop, "Female Earnings Inequality: The Changing Role of Family Characteristics and Its Effect on the Extensive and Intensive Margins," *Journal of Labor Economics*, 2021, *39* (S1), S59–S106.

_ , Jesse Rothstein, and Moises Yi, "Industry Wage Differentials: A Firm-Based Approach," *Journal of Labor Economics*, 2024, *42* (S1), 11–59. Earlier version: NBER Working Paper No. w31588 (DOI: 10.3386/w29920).

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

Cardoso, Ana Rute, Paulo Guimarães, and Pedro Portugal, "What Drives the Gender Wage Gap? A Look at the Role of Firm and Job-Title Heterogeneity," *Oxford Economic Papers*, jan 2016, *68* (2), 506–524.

Casarico, Alessandra and Salvatore Lattanzio, "What Firms Do: Gender Inequality in Linked Employer-Employee Data," *Journal of Labor Economics*, April 2024, *42* (2), 325–355.

Chikhale, Nisha, Natalie Duncombe, and Birthe Larsen, "The Scarring Effects of Workplace Sexual Harassment," WorkingPaper 07-2025, Department of Economics. Copenhagen Business School 2025.

Coudin, Elise, Sophie Maillard, and Maxime Tô, “Family, Firms and the Gender Wage Gap in France,” IFS Working Papers W18/01, Institute for Fiscal Studies 2018. Accessed: 19 August 2024.

Duncan, Otis Dudley and Beverly Duncan, “A Methodological Analysis of Segregation Indexes,” *American Sociological Review*, April 1955, *20* (2), 210.

Farber, Henry S., “Chapter 37 Mobility and stability: The dynamics of job change in labor markets,” in “in,” Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, pp. 2439–2483.

Folke, Olle and Johanna Rickne, “Sexual Harassment and Gender Inequality in the Labor Market,” *The Quarterly Journal of Economics*, may 2022, *137* (4), 2163–2212.

Gallen, Yana, Rune V. Lesner, and Rune Vejlin, “The labor market gender gap in Denmark: Sorting out the past 30 years,” *Labour Economics*, 2019, *56*, 58–67.

Goldin, Claudia, “A Grand Gender Convergence: Its Last Chapter,” *American Economic Review*, 2014, *104* (4), 1091–1119.

_ , Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth, “The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census,” *American Economic Review*, May 2017, *107* (5), 110–114.

Guimarães, Paulo and Pedro Portugal, “A Simple Feasible Procedure to fit Models with High-dimensional Fixed Effects,” *The Stata Journal: Promoting communications on statistics and Stata*, January 2011, *10* (4), 628–649.

Haltiwanger, John C., Henry R. Hyatt, and James Spletzer, “Industries, Mega Firms, and Increasing Inequality,” NBER Working Paper 29920, National Bureau of Economic Research April 2022. Revised December 2022.

Hotz, V. Joseph, Per Johansson, and Arizo Karimi, “Parenthood, Family Friendly Workplaces, and the Gender Gaps in Early Work Careers,” NBER Working Paper 24173, National Bureau of Economic Research 2018.

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

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante, “The Child Penalty Atlas,” *The Review of Economic Studies*, 10 2024, *92* (5), 3174–3207.

_ , _ , and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 181–209.

Labanca, Claudio and Dario Pozzoli, "Constraints on Hours within the Firm," *Journal of Labor Economics*, April 2022, *40* (2), 507–539.

Lhuillier, Hugo, "The Local Root of Wage Inequality," BFI Working Paper 2025-78, Becker Friedman Institute for Economics, University of Chicago June 2025. Available at SSRN.

Maasoumi, Esfandiar and Le Wang, "The Gender Gap between Earnings Distributions," *Journal of Political Economy*, 2019, *127* (5), 2438–2504.

Manning, Alan and Barbara Petrongolo, "The part-time pay penalty for women in Britain," *The Economic Journal*, 2008, *118* (526), F28–F51.

Masso, Jaan, Jaanika Meriküll, and Priit Vahter, "The role of firms in the gender wage gap," *Journal of Comparative Economics*, June 2022, *50* (2), 454–473.

Morchio, Iacopo and Christian Moser, "The Gender Pay Gap: Micro Sources and Macro Consequences," NBER Working Paper 32408, National Bureau of Economic Research May 2024.

Palladino, Marco G., Alexandra Roulet, and Mark Stabile, "Narrowing industry wage premiums and the decline in the gender wage gap," *Labour Economics*, June 2025, *94*, 102693.

Petrongolo, Barbara and Maddalena Ronchi, "Gender gaps and the structure of local labor markets," *Labour Economics*, June 2020, *64*, 101819.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality*," *The Quarterly Journal of Economics*, October 2018, *134* (1), 1–50.

Sorkin, Isaac, "The Role of Firms in Gender Earnings Inequality: Evidence from the United States," *American Economic Review*, 2017, *107* (5), 384–387.

_ , "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, August 2018, *133* (3), 1331–1393.

Targa, Matteo, "Are 'good' firms, good for all employees?," *Labour Economics*, December 2025, *97*, 102816.

Torres, Sónia, Pedro Portugal, John T. Addison, and Paulo Guimarães, "The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model," *Labour Economics*, October 2018, *54*, 47–60.

Woodcock, Simon D., "Wage differentials in the presence of unobserved worker, firm, and match heterogeneity," *Labour Economics*, August 2008, *15* (4), 771–793.

Online Appendix for "Hidden in Plain Sight: Occupational Structure and the Gender Wage Gap"

Riccarda Rosenball
University of Graz

Tobias Eibinger
University of Graz

Jörn Kleinert
University of Graz

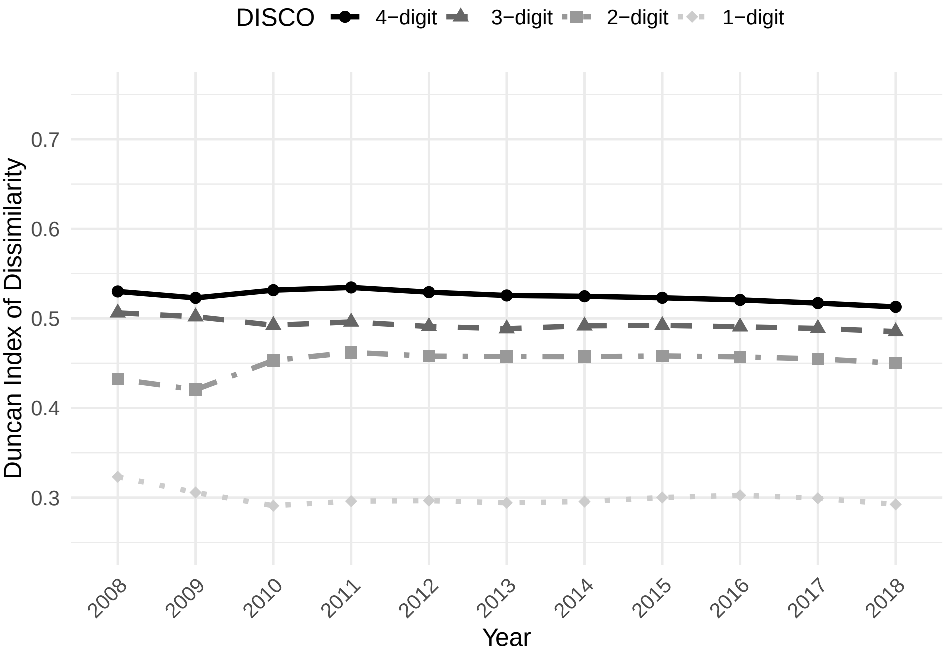
Ismir Mulalic
Copenhagen Business School

January 29, 2026

Table of Contents

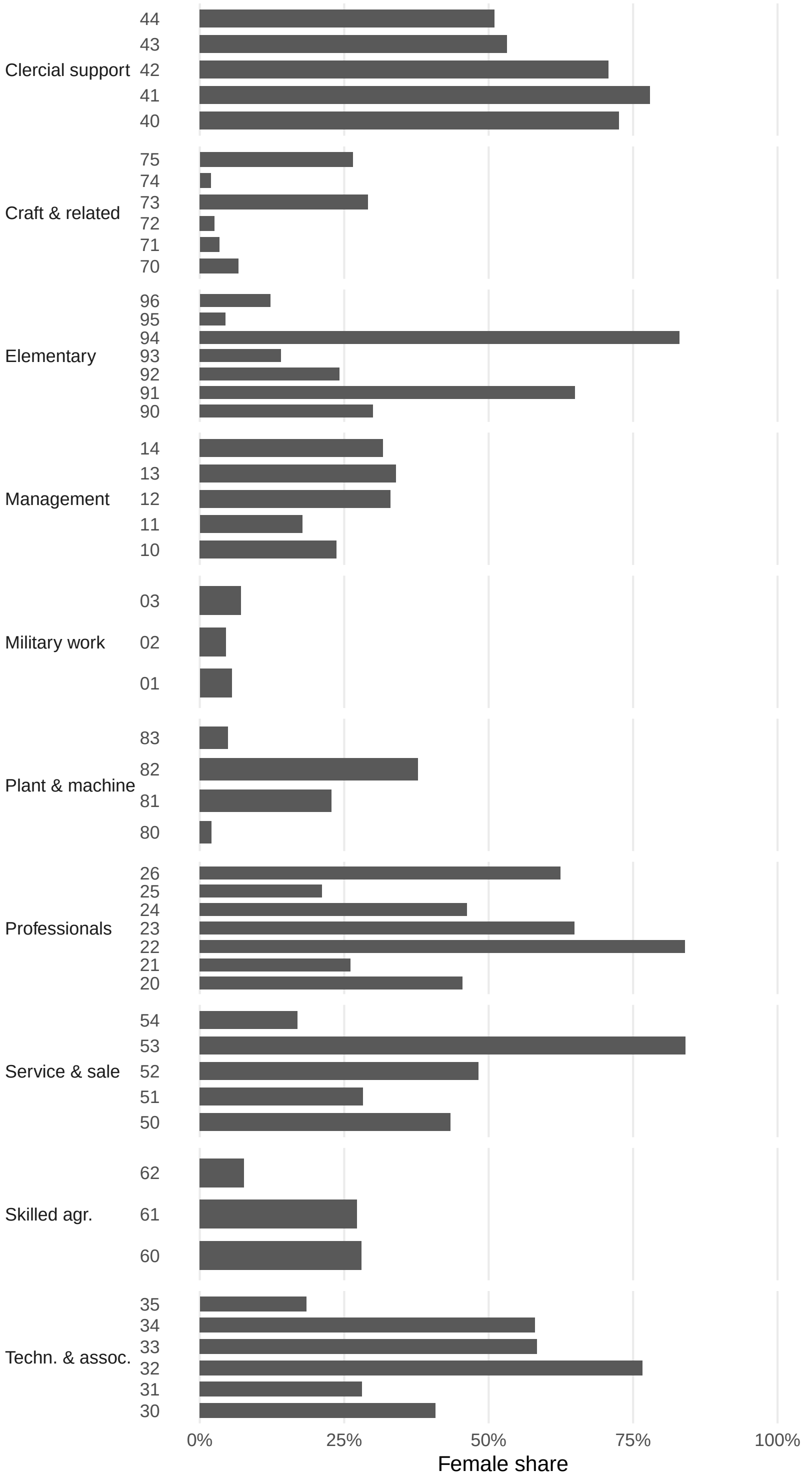
A Additional Data Descriptives

Figure A.1. Duncan index of gender segregation across occupations over time



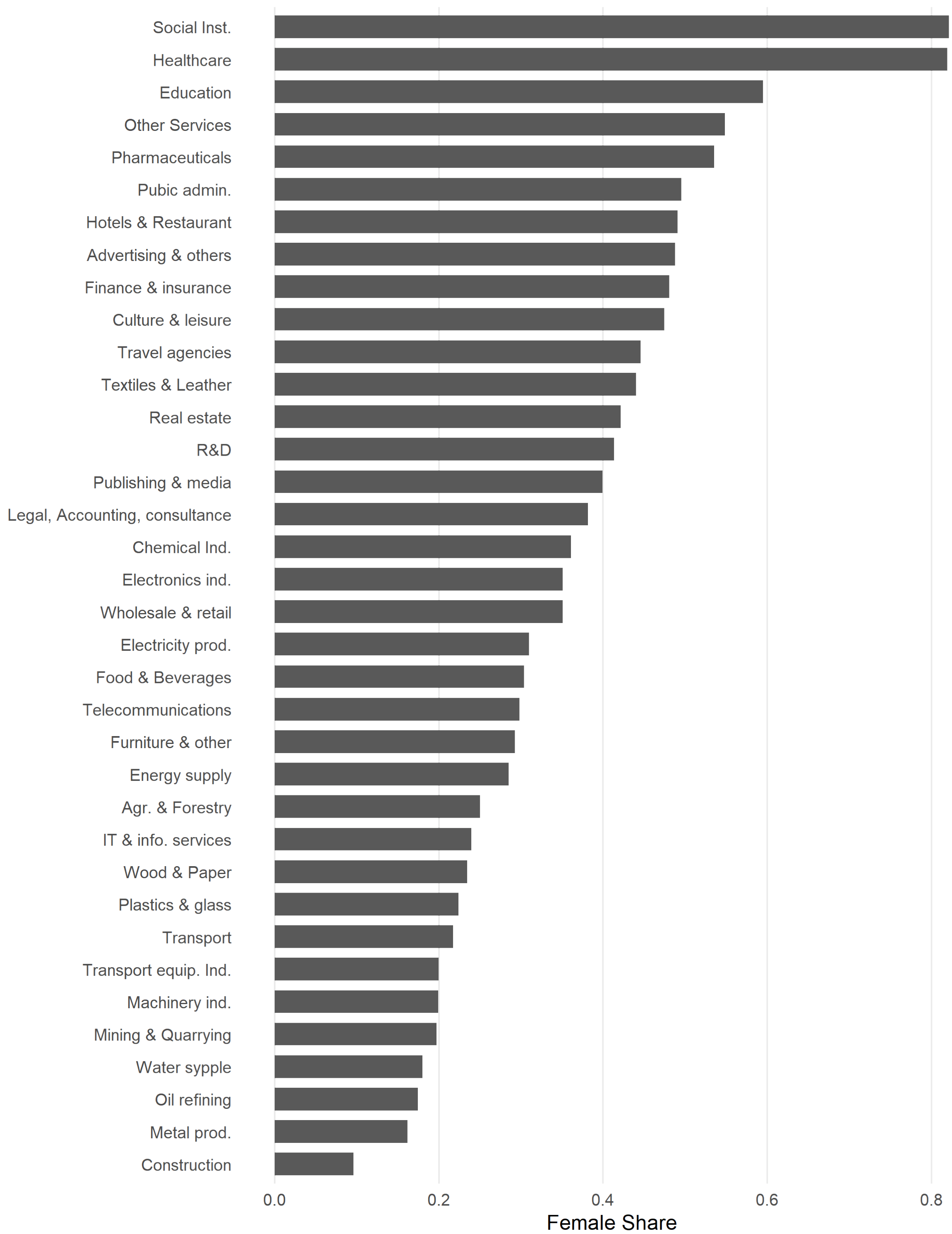
Note: The figure plots the Duncan index values over time for different levels of occupational aggregation (DISCO 1–4 digits). The Duncan index of dissimilarity is a statistical measure of segregation that shows how evenly two groups are distributed across different areas. It ranges from 0 (complete integration) to 1 (complete segregation) and indicates the proportion of one group that would need to move for the groups to be evenly distributed.

Figure A.2. Female share by 2-digit DISCO (2013–2018)



Note: The figure plots the female share across 2-digit occupations, grouped into the following 1-digit occupational categories: clerical support work, craft and related trades, elementary occupations, management occupations, military work, plant and machine operators, professionals, service and sales work, skilled agricultural work, and technicians and associate professionals.

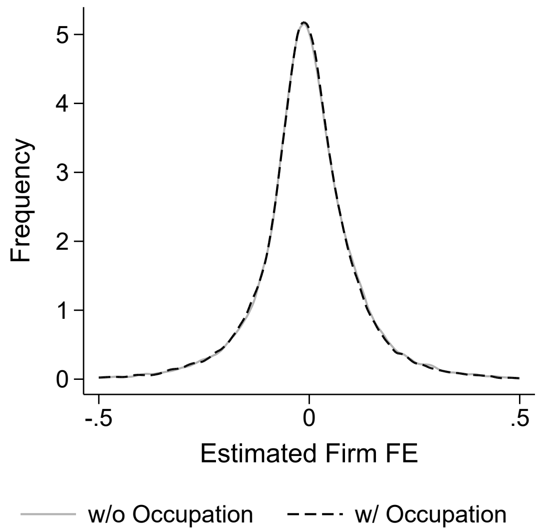
Figure A.3. Female share by Industry (2013 - 2018)



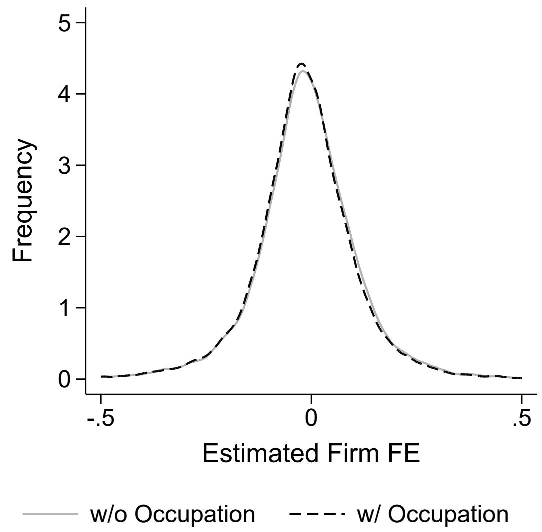
Note: The figure shows the female share across 36 sub-industries, ordered from the most to the least female-dominated industry.

B Additional Results

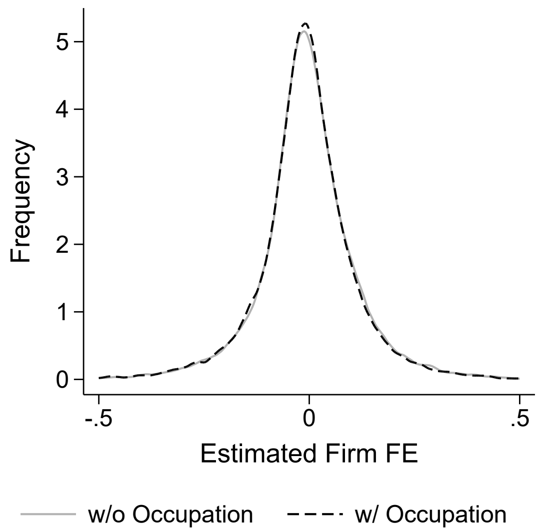
Figure B.1. Kernel densities of firm fixed effects estimated with AKM, with and without occupation controls (2013–2018).



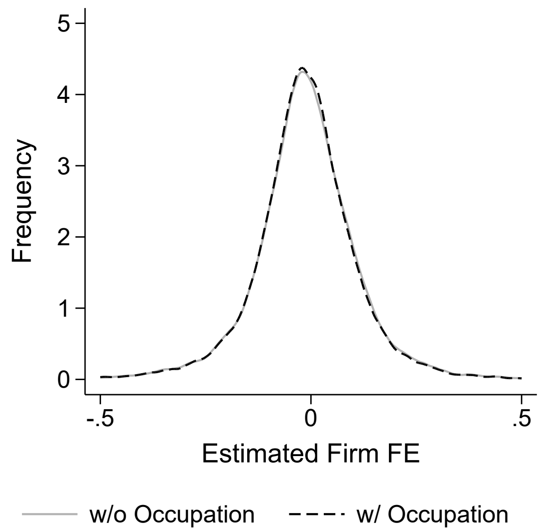
(a) Women, 2-digit DISCO



(b) Men, 2-digit DISCO



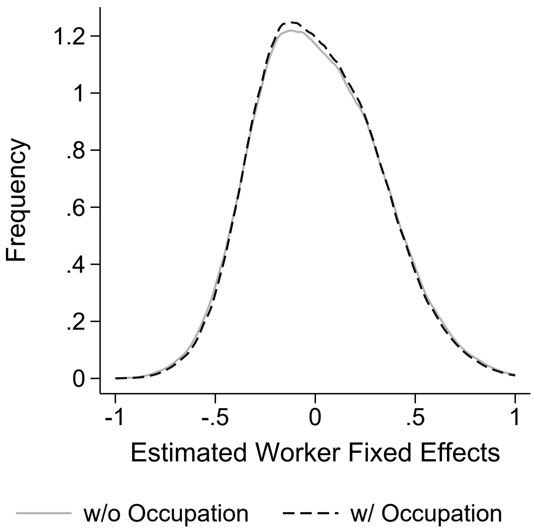
(c) Women, 4-digit DISCO



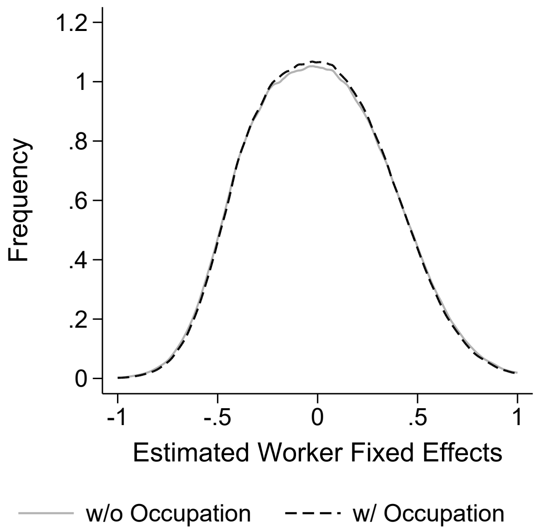
(d) Men, 4-digit DISCO

Note: The left panels display kernel density estimates of firm fixed effects from the male AKM regression, comparing specifications without occupation controls (dashed line) and with occupation controls (solid line). The right panels show the corresponding densities from the female AKM regression. The upper panels include occupation controls at the 2-digit level, while the lower panels use more detailed 4-digit occupation controls.

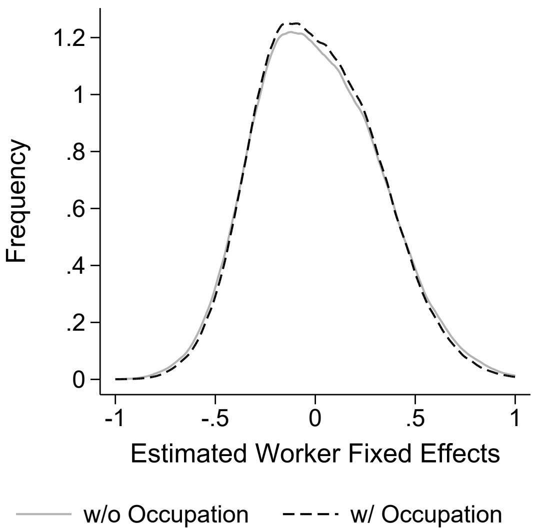
Figure B.2. Kernel densities of worker fixed effects estimated with AKM, with and without occupation controls (2013–2018).



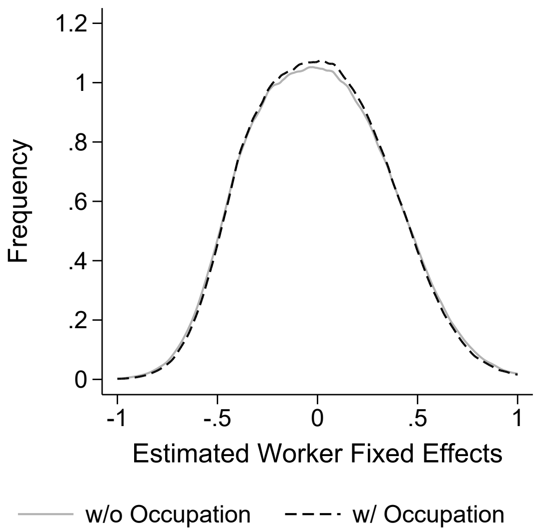
(a) Women, 2-digit DISCO



(b) Men, 2-digit DISCO



(c) Women, 4-digit DISCO



(d) Men, 4-digit DISCO

Note: The left panels display kernel density estimates of worker fixed effects from the male AKM regression, comparing specifications without occupation controls (dashed line) and with occupation controls (solid line). The right panels show the corresponding densities from the female AKM regression. The upper panels include occupation controls at the 2-digit level, while the lower panels use more detailed 4-digit occupation controls.

Table B.1. Decomposition of the gender wage gap (GWG) by period and specification: Male firm fixed effects reference

Specification	DISCO	Raw GWG	Sorting	Bargaining	Firm	Firm/GWG	Sorting/GWG
Panel A: 2008–2012							
W/o occupation	–	0.1683	0.0213	0.0032	0.0245	14.6%	12.7%
W/ occupation	1-digit	0.1683	0.0206	0.0033	0.0239	14.2%	12.2%
W/ occupation	2-digit	0.1683	0.0200	0.0029	0.0229	13.6%	11.9%
W/ occupation	3-digit	0.1683	0.0193	0.0019	0.0212	12.6%	11.5%
W/ occupation	4-digit	0.1683	0.0188	0.0050	0.0238	14.2%	11.2%
Panel B: 2013–2018							
W/o occupation	–	0.1402	0.0212	−0.0086	0.0125	8.9%	15.1%
W/ occupation	1-digit	0.1402	0.0199	−0.0080	0.0119	8.5%	14.2%
W/ occupation	2-digit	0.1402	0.0179	−0.0115	0.0065	4.6%	12.8%
W/ occupation	3-digit	0.1402	0.0178	−0.0092	0.0086	6.1%	12.7%
W/ occupation	4-digit	0.1402	0.0176	−0.0080	0.0096	6.9%	12.6%

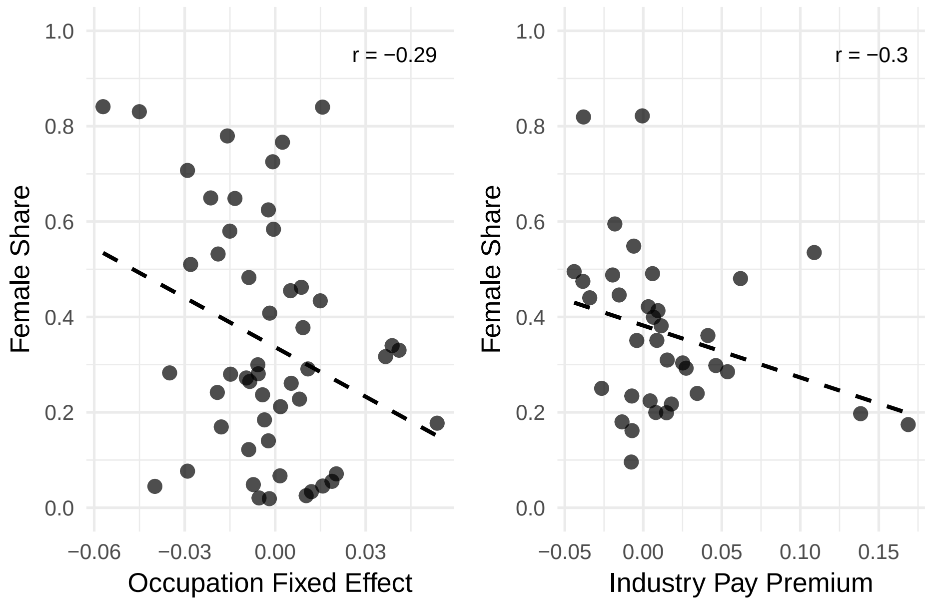
Note: Decomposition of the gender wage gap into sorting and bargaining components using female firm fixed effects as the reference group. Estimates are reported for two periods (2008–2012 and 2013–2018) and for AKM specifications with increasing levels of occupational detail. "Firm" denotes the firm-specific component of the decomposition, and the percentages express each component relative to the raw gender wage gap.

Table B.2. Decomposition of firm sorting into between- and within-industry components (male firm fixed effects reference)

Specification	DISCO	Raw GWG	Sorting	Between ind.	Within ind.	Share btw.	Share within
Panel A: 2008–2012							
w/o occupation	-	0.1683	0.0213	0.0190	0.0023	89.1%	10.9%
w/ occupation	1-digit	0.1683	0.0206	0.0184	0.0022	89.3%	10.7%
w/ occupation	2-digit	0.1683	0.0200	0.0178	0.0022	89.1%	10.9%
w/ occupation	3-digit	0.1683	0.0193	0.0171	0.0022	88.6%	11.4%
w/ occupation	4-digit	0.1683	0.0188	0.0166	0.0023	88.0%	12.0%
Panel B: 2013–2018							
w/o occupation	-	0.1402	0.0212	0.0193	0.0018	91.3%	8.7%
w/ occupation	1-digit	0.1402	0.0199	0.0183	0.0016	91.9%	8.1%
w/ occupation	2-digit	0.1402	0.0179	0.0162	0.0017	90.7%	9.3%
w/ occupation	3-digit	0.1402	0.0178	0.0162	0.0015	91.3%	8.7%
w/ occupation	4-digit	0.1402	0.0176	0.0159	0.0017	90.4%	9.6%

Note: Decomposition of the firm sorting component of the gender wage gap into between- and within-industry sorting based on 128 industry categories. The between-industry component captures gender differences in the allocation of workers across industries, while the within-industry component reflects sorting across firms within the same industry. Estimates use *male firm fixed effects* as the reference. The percentages report the share of between-industry (btw.) and within-industry (within) sorting in total firm sorting.

Figure B.3. Correlations between female share, occupation fixed effects, and industry pay premia



Note: The left panel presents the correlation between the female share of an occupation with the occupation fixed effect for occupations on the 2-digit level for the period 2013 to 2018. The right panel shows the correlations between the female share of the industry (36 sub-industries) and the industry pay premium.

From figure B.3 in Appendix B we also absorve a pronounced negative correlation of about -0.3 between the female share across 2 digit occupations the the occupation fixed effect and between the female share across 36 sub-industries and the industry pay premium. Indicating women are increasingly represented at lower paying occpuations and industries.

Table B.3. Occupations, DISCO Codes, Male Occupation Fixed Effects, and Ranks

Occupation	DISCO	Occ. FE (Male)	Rank
Care work	53	−0.057073485	1
Manual food prepartion work	94	−0.045041812	2
Street sales and services	95	−0.039893709	3
Service work	51	−0.034996175	4
Customer services	42	−0.029085996	5
Work in forestry, fishing and hunting	62	−0.029070924	6
Other general office and customer service work	44	−0.02804033	7
Cleaning work	91	−0.021380079	8
Manual work in agriculture, forestry, and fishing	92	−0.019190383	9
Gernal calculation and registration work	43	−0.018930713	10
Rescue and surveillance work	54	−0.017883218	11
General office and secretarial work	41	−0.015868462	12
Work in law, social sciences and cultrure	34	−0.015050841	13
Skilled agricultural, forestry and fishery workers	60	−0.014804061	14
Teaching and educational work	23	−0.013347178	15
Work in agriculture and horticulture	61	−0.009584725	16
Waste collection and other manualwork	96	−0.008781388	17
Sales work	52	−0.008718701	18
Food production, carpentry, clothing production, and related crafts	75	−0.008434499	19
Driving work and operators of vehicles and mobile machinery	83	−0.00725744	20
Elementary occupations	90	−0.005752608	21

Occupation	DISCO	Occ. FE (Male)	Rank
Technician work in science, engineering, shipping and aviation	31	−0.005627291	22
Plant and machine operators and assemblers	80	−0.005363871	23
Managers (broad)	10	−0.004202741	24
Information and communication technician work	35	−0.003557783	25
Manual work in mining, construction, production and transport	93	−0.002271797	26
Work in law, scial sciences, and culture	26	−0.002269607	27
Work in electrical and electronic fields	74	−0.001937513	28
Technicians and associate (broad)	30	−0.001831307	29
Clercial support workers (broad)	40	−0.000829192	30
Work in business services, economics, administration and sales	33	−0.000575983	31
Craft and related trades workers	70	0.001592477	32
Work in information and communication technology	25	0.001745451	33
Technician and assistant work in health	32	0.002404365	34
Pressionals (broad)	20	0.005082279	35
Work in naturalsciences and engineering	21	0.005307377	36
Operator work at fixed installations and machinery	81	0.008040475	37
Work in economics, administration and sales	24	0.008662302	38
Assembly work	82	0.00920821	39
Metal and machinery work	72	0.01027104	40
Precision craft and graphic work	73	0.010839519	41
Craft-based work in construction, excluding electrical work	71	0.012002854	42
Sercie and sales workers (broad)	50	0.014936861	43
Work in health sector	22	0.015689406	44

Occupation	DISCO	Occ. FE (Male)	Rank
Military work at non-commissioned officer level	2	0.015767234	45
Military work at officer level	1	0.018821561	46
Military work, other ranks	3	0.020263723	47
Management in hotel & restaurant, retail and wholesale trade and other services	14	0.036588231	48
Management of core activities in production and service enterprises	13	0.038764648	49
Management in administration and business-oriented functions	12	0.041063459	50
Senior management in legislative bodies businesses, and organizations	11	0.053687719	51

Table B.4. Industry Short Codes, Pay Premia (with occupation controls), and Ranks

Industry	Short	Pay Premium	Rank
Public administration	PA	−0.044043429	1
Culture and leisure	CL	−0.038472757	2
Healthcare	HC	−0.038081728	3
Textiles and Leather Industry	TL	−0.034103498	4
Agriculture and Forestry	AF	−0.026514698	5
Advertising and other business services	AB	−0.019528730	6
Education	E	−0.018110223	7
Travel agencies	TA	−0.015356165	8
Water supply	WS	−0.013545454	9
Construction	CO	−0.007644010	10
Wood and Paper	WP	−0.007308495	11
Metal production	MP	−0.007168583	12
Other Services	OS	−0.006085531	13
Wholesale and retail trade	WR	−0.004142614	14
Social Institutions	SI	−0.000625487	15
Real estate activities	RE	0.003218659	16
Plastics and glass manufacturing industry	PG	0.004330726	17
Hotels and Restaurant	HR	0.005893402	18
Publishing, TV and media	PTV	0.006479328	19
Transport equipment industry	TE	0.007967528	20
Electronics industry	EL	0.008655259	21
Research and Development	RD	0.009364107	22
Legal, Accounting, and consultancy services	LAC	0.011394151	23
Machinery industry	MI	0.014811900	24
Electricity production	ET	0.015240405	25
Transport	T	0.017930632	26
Food and Beverages	FB	0.025112020	27
Furniture and other manufacturing	FM	0.027287588	28
IT and information services	IT	0.034345377	29
Chemical Industry	CH	0.041157413	30
Telecommunications	TC	0.046214301	31
Energy supply	EG	0.053651154	32
Financial and insurance activities	FI	0.061934572	33
Manufacture of pharmaceutical products	PH	0.108894072	34

Industry	Short	Pay Premium	Rank
Mining and Quarrying	MQ	0.138430908	35
Oil refining	OR	0.168749809	36