Worker Sorting and the Gender Wage Gap

Giulia Vattuone*

March 26, 2026

Abstract

Part of the gender wage gap reflects women sorting into lower-paying firms. Using French matched employer-employee data, this paper examines whether gender differences in firm allocation reflect preferences or opportunities, and how they vary across the wage distribution and over the career. A finite-mixture model recovers latent worker types and firm classes from wages, mobility, and observables, allowing wage and mobility profiles to vary by gender and career stage within type. Counterfactuals show that sorting matters at the top. Among high-wage workers, equalising women's mobility to men's reduces the lifecycle wage gap by 29 percent. Preferences over firms explain more than half of this sorting effect, while re-entry after non-employment explains most of the remainder. Equalising wage growth and mobility closes 80 percent of the high-wage lifecycle gap, indicating slower wage progression for women even conditional on worker type, firm class, career stage, and occupation.

JEL Codes: J16; J31; J63; J64

Keywords: Gender Wage Gap; Sorting; Job Preferences; Job Mobility; Finite Mixture Models

* SOFI, Stockholm University. Email: giulia.vattuone@sofi.su.se. I thank Roland Rathelot and Manuel Bagues for their guidance and support. I am especially grateful to Suphanit Piyapromdee and Rune Vejlin for helpful comments. For insightful conversations and feedback, I thank Christian Belzil, Christine Braun, Luis Candelaria, Julien Combe, Oren Danieli, Mirko Draca, Emma Duchini, James Fenske, Rasmus Lentz, Erik Lindqvist, Joan Llull, Jonas Maibom, Roberto Pancrazi, Eric Renault, Seth Sanders, Benoit Schmutz, Isaac Sorkin, Thijs van Rens, Natalia Zinovyeva and seminar participants at Warwick University, CREST-ENSAE, Aarhus University, and Stockholm University. I would also like to thank Christian Hellwig, Olivier Scaillet, and participants at the EEA-ESEM 2024 Congress. An earlier version of this paper was awarded the ESEM 2024 Award for Best Paper in Applied Economics. I thank CASD for providing data access and support. All errors are my own.

1 Introduction

The gender wage gap partly reflects differences in sorting across firms. Following the seminal work of Card, Cardoso, and Kline (2016), several studies confirm this finding.[1] A major debate is whether gender differences in sorting stem from differences in preferences or opportunities. This paper addresses this question by estimating a model of wages and mobility by exploiting information on firm-to-firm transitions.

A long-standing literature shows that gender differences in mobility are not primarily about whether workers move but rather about the wage gains associated with moving (Loprest, 1992). Differential wage gains may arise because some firms pay higher wages to observably similar workers. The gender wage gap widens over the career, especially in the early years, and part of this divergence comes from differential sorting across firms (Goldin, Kerr, Olivetti, and Barth, 2017; Barth, Kerr, and Olivetti, 2021). The question is therefore whether differences in firm allocation over the career are driven by the employment opportunities women face or by the firms they choose conditional on those opportunities, and how the relative importance of these channels varies across segments of the wage distribution.

To answer this question, I employ a revealed preference argument in a random search framework, initially proposed by Sorkin (2018). The basic idea is that observed firm-to-firm transitions are informative about offer arrival rates and worker preferences. Upon receiving an offer, a worker chooses to accept if the perceived value of the poacher is higher than the one of the incumbent. Workers may value something beyond wages in a way that guides where they sort. Throughout the paper, offer arrival rates represent employment opportunities to

1. It has been widely documented that unequal gender distributions across workplaces contribute to the gender wage gap (Blau, 1977; Hirschman, 2022). Card, Cardoso, and Kline (2016) is the first paper to comprehensively analyse the role of gender differences in worker-firm allocations in explaining the gender wage gap. Their approach, which builds on the log earnings model of Abowd, Kramarz, and Margolis (1999), has been adopted using data from multiple countries: the US (Sorkin, 2017), France (Coudin, Maillard, and To, 2018; Palladino, Roulet, and Stabile, 2025c), Germany (Bruns, 2019), Italy (Casarico and Lattanzio, 2024), Canada (Li, Dostie, and Simard-Duplain, 2020), Brazil (Morchio and Moser, 2023), Chile (Cruz and Rau, 2022), and in cross-country analyses (Palladino et al., 2025b). The share of the gender wage gap due to differences in firm sorting ranges roughly between 15% and 25%, depending on country-specific data availability and labour market institutions. Differences in firm sorting are not related to a lack of skills. In general, Blau and Kahn (2017) stresses that conventional supply-side factors, such as human capital accumulation, psychological attributes, and non-cognitive skills, cannot explain a substantial portion of the gender wage gap. See also Olivetti and Petrongolo (2016) for an extensive literature review of gender gaps.

move to a specific firm, while perceived firm values represent worker preferences.

I combine the revealed preference approach with a finite-mixture model, recently proposed by Lentz, Piyapromdee, and Robin (2023) and Bonhomme, Lamadon, and Manresa (2019), in order to allow for worker heterogeneity within and between genders in a framework that generates rich sorting patterns across different career stages. I estimate the model on matched employer-employee quarterly data for the region *Île-de-France* (Greater Paris) over 2015–2019.

Administrative data directly provide realised worker-firm matches, making it difficult to disentangle choices from opportunities to move. I address the problem with a finite-mixture model that groups workers into latent types and firms into latent classes, allowing these latent groups to vary flexibly with gender and career stage. A latent representation of heterogeneity enables separating opportunities from choices while preserving the heterogeneity needed to study sorting over the career and across segments of the wage distribution. The key assumption is that latent groups provide a tractable approximation to the dimensions of worker and firm heterogeneity that matter for the joint distribution of wages and mobility.

Firm-to-firm transition probabilities are modelled as the product of an offer arrival rate and a choice probability. The decision to move depends on perceived firm values as well as an incumbent wedge, which may raise the value of staying with the current employer and allow for systematic differences in reluctance to move across worker types and career stages. Identification relies on within-class transitions and asymmetries in bilateral flows between firm classes, which help separately identify offer arrival rates and workers' preferences over firms. Transitions into non-employment and out of non-employment are left unrestricted.

Finally, the framework allows for worker-firm wage complementarities, assuming that workers draw hourly wages from a distribution specific to worker types and firm classes. Similar to Abowd, Kramarz, and Margolis (1999), mobility depends only on worker types and firm classes but not directly on wages. Unlike that literature, however, wage distributions are estimated using information from both movers and stayers. Borovickova and Shimer (2025) emphasize that wages and firm pay policies inferred from movers alone may be based on a non-representative sample of wages. Although the model of this paper maintains the exogenous mobility assumption, using data on both movers and stayers should attenuate that concern.

I estimate the model in two steps, following Bonhomme, Lamadon, and Manresa (2019). First, I cluster firms into classes employing a k-means algorithm. This algorithm uses firm data on size, gender-specific wage distributions, and female shares. Second, conditional on the firm classes, I cluster workers into types and estimate the parameters of interest using an Expectation-Maximisation algorithm, which uses data on observed workers' wages, characteristics, and transitions between firm classes and employment statuses. Latent worker types interact flexibly with gender, age, and tenure, allowing wage profiles and mobility patterns of men and women of the same type to vary over the career. I adapt the iterative procedure of Lentz, Piyapromdee, and Robin (2023) to handle the non-linear mobility parameters and the incumbent wedge. Wages are residualized on 3-digit occupation and year fixed effects, and worker types are ordered from low-wage to high-wage, and firm classes from low-paying to high-paying, through a standard relabelling of the estimated clusters.

The clustering results reveal that worker types capture distinct career paths. Low-wage types exhibit nearly flat wage profiles, whereas mid- and especially high-wage types display steep wage growth over the career, with the gender wage gap widening most strongly at the top, echoing results from Goldin (2014) and Goldin, Kerr, Olivetti, and Barth (2017).

Preference heterogeneity is also substantial. Estimated rankings of firm classes differ not only across worker types within gender but also across men and women of the same type, with the largest cross-gender divergence arising early in the career. Using linked *Conditions de Travail* survey data, which provide suggestive evidence given limited firm coverage, I find a sharp gender divergence among high-wage workers at early career stages: conditional on pay, women's preferences are strongly negatively correlated with longer hours and overtime, unlike men's. Longer hours and overtime are also more prevalent in higher-paying firms. Differences in valuations can therefore translate into different acceptance decisions and, in turn, into different firm allocations that may cumulate into larger gender wage gaps over the life cycle.

To quantify the role of sorting, I simulate cohorts of workers from age 25 and follow them quarter by quarter for twenty years under the estimated model. Starting from the estimated distribution of initial worker-firm matches, I compare baseline wage profiles with a counterfactual in which women face the full set of male mobility parameters while retaining their own wage schedule in each worker-firm match, so that any change in wages comes from

reallocation across firms and employment states rather than from changes in pay conditional on a given match.

Equalising mobility raises women's wages in all worker types, but the effect is highly uneven across the wage distribution. The integrated lifecycle wage gap falls by little at the bottom, by 11% among mid-wage workers, and by 29% among high-wage workers. Because the lifecycle gap is both steepest and most responsive to the counterfactual at the top, I focus the main decomposition on high-wage workers.[2]

To quantify the relative importance of multiple mobility channels, I perform several counterfactual exercises where I equate the mobility parameters of women to the ones of men to simulate scenarios where men and women progressively share similar *i*) offer arrival rates while in employment, *ii*) preferences over firms, *iii*) incumbent wedge, *iv*) transitions into non-employment, and *v*) offer arrival rates while in non-employment. I study how the gender wage gap changes under these multiple scenarios.

First, among high-wage workers, gender differences in the job offer distribution while in employment do not contribute to the gender wage gap. On the contrary, if women sampled job offers at the same frequency as men, the gender wage gap would increase, implying that women are more likely to draw job offers that would pay more. Although surprising, this result may be supported by the findings of a recent correspondence study run by the French *Institut des Politiques Publiques*. The study carried out a large-scale experiment by sending fictitious CVs in response to several thousand job offers in eleven distinct professions. Callback rates in low-skilled occupations are significantly lower for women. In contrast, the opposite is observed for executive occupations with supervision, roles populated mainly by high-experience men.[3] To the extent that callbacks reflect actual job offers and that the firm clustering captures differences in occupational compositions, my results align with these findings.

Second, the main gap reduction begins once women face male firm valuations, which account for more than half of the sorting effect. Among high-wage workers, this channel is

2. I report results for other types in the Appendix.

3. (Note IPP n°67) Discrimination à l'embauche selon le sexe: les enseignements d'un testing de grande ampleur. https://www.ipp.eu/actualites/note-ipp-n67-discrimination-a-l-embauche-selon-le-sexe-les-enseignements-d-un-testing-de-grande-ampleur/

especially important at mid-career stages, where the baseline female wage profile displays a marked flattening around age 35. When women instead value firms as men do, that flattening is substantially attenuated and the lifecycle gap opens more gradually. Preferences over firms therefore emerge as the main job-to-job margin through which sorting contributes to the widening gender wage gap at the top.

Third, equalising the incumbent wedge and exit rates contributes little further. I find strong cross-gender correlations in the estimates of these parameters between female and male workers.

Fourth, I find that gender differences in the offer arrival rates while in non-employment are the second most important determinant of the sorting effect. While some aspects of the preference mechanism will almost certainly influence re-entry patterns, I find extremely weak correlations between female and male entry rates across *all* worker types, also among those with a strong alignment in the preference parameter. Following periods of non-employment, women are more likely to draw job offers that pay them less.

The role of sorting appears to vary over the career. The compression in the high-wage lifecycle gap is strongest around the mid-thirties and more limited later in the career, suggesting that reallocation across firms becomes less central as careers unfold. I therefore add a final counterfactual that equalises women's residual wage growth to the male profile after equalising mobility. Under this exercise, women retain their own wage level at labor-market entry within each tenure group, but inherit the male wage-growth increment over the age-tenure dimension. The resulting compression is large: among high-wage workers at age 45, the lifecycle gap falls from 17.9 log points at baseline to 14.5 log points under equal mobility, and to 3.7 log points once residual wage growth is also equalised. In cumulative terms, about 80% of the high-wage lifecycle gap disappears. The result therefore points to slower wage progression for women even conditional on worker type, firm class, age-tenure stage, and common 3-digit occupation and year effects, while the remaining 20% of the high-wage lifecycle gap reflects persistent residual differences that the model does not explain.

This paper relates to several strands of literature. First, long-standing literature has been studying gender differences in labour mobility in determining wage differentials (Loprest, 1992; Bowlus, 1997; Del Bono and Vuri, 2011). Compared to this literature, I leverage detailed

matched employer-employee data.

Second, the paper speaks to the literature that quantifies the sorting effect of the gender wage gap. This literature starts with Card, Cardoso, and Kline (2016) and builds on the pioneering work of Abowd, Kramarz, and Margolis (1999), who estimate by Ordinary Least Squares a linear wage equation with additive worker and firm fixed effects and condition on observed worker characteristics. Adopting the finite mixture model of Lentz, Piyapromdee, and Robin (2023) and Bonhomme, Lamadon, and Manresa (2019) permits the explicit modelling of mobility, allowing me to gauge the relative importance of key mobility channels driving the sorting effect. Casarico and Lattanzio (2024) show that women are less likely to move towards firms with higher wage policies upon firm-to-firm moves. My paper complements their result by separating the role of offer arrival rates and worker preferences in firm-to-firm transitions.

Most importantly, estimating worker-perceived firm values connects my paper to Sorkin (2017) and Morchio and Moser (2023). Sorkin (2017) uses a revealed-preference approach and finds that men and women value firms similarly, attributing between-firm pay gaps to differences in job offer distributions. Morchio and Moser (2023) develop an equilibrium model where firms' optimal recruiting decisions identify firm-level utility offers and find that compensating differentials explain half of the gender gap in firm wage policies. The contribution of my paper is to study sorting over the lifecycle while preserving rich sources of worker heterogeneity. I do not assume homogenous preferences but allow worker types to have their own preferences, which evolve with tenure and tenure. These sources of worker heterogeneity are valuable as within-gender variation may swamp average gender differences in some mobility factors in a way that may underestimate the relevance of gender differences in sorting across different market segments and at different career stages.[4] I find substantial dispersion in preferences across worker types.

4. For example, Bertrand (2020) stresses the importance of within-gender variation in personal traits such as confidence, risk aversion, and willingness to negotiate. She reviews several meta-analyses that conclude that average gender gaps in these personal traits are minimal. In my framework, allowing for worker heterogeneity within gender is empirically important. In a model with one worker type, within-gender disagreement is mechanically absent and cross-gender rank divergence is more muted. Once worker heterogeneity is introduced, substantial disagreement emerges both within gender and across genders, with the largest cross-gender divergence appearing early in the career.

More broadly, the paper contributes to a literature that emphasizes the central role of heterogeneity in shaping economic behaviour (Heckman and Singer, 1984; Bonhomme and Denis, 2025; Samuelson and Steiner, 2025). In that spirit, the latent-type approach is not only a computational device but a way to capture economically meaningful variation in wage and mobility responses within and across gender over the life cycle.

Finally, I relate to the important literature that points out that gender wage differentials may materialize as a result of differences in job search behaviour (Cortés, Pan, Pilossoph, Reuben, and Zafar, 2023; Braun and Figueiredo, 2025), as a result of employer discrimination in hiring (Neumark, Bank, and Nort, 1996; Xiao, 2023; Kline, Rose, and Walters, 2022), or as women have stronger preferences for shorter commuting time (Petrongolo and Ronchi, 2020; Le Barbanchon, Rathelot, and Roulet, 2021; Caldwell and Danieli, 2024; Fluchtmann, Glenny, Harmon, and Maibom, 2024), or for flexibility (Wiswall and Zafar, 2018).[5] In this paper, I attempt to separate the relative importance of gender differences in offer distributions, which subsume worker and firm search behaviour, and in preferences. I infer worker preferences from firm-to-firm transitions and do not focus on a specific preference mechanism. My estimates of worker-perceived firm values capture an overall bundle of firm characteristics valued by workers.[6] Throughout the paper, I do not take a stand on whether gender differences in perceived firm values arise from 'true' preferences or whether they reflect gender stereotypes or norms that influence the choices men and women make.

The remainder of the article is organized as follows. Section 2 presents the framework of analysis. Section 3 describes the estimation procedure. Section 4 and Section 5 present the data and results from the classification algorithms. Section 6 illustrates the counterfactual analysis. Finally, section 7 concludes and discusses some caveats.

5. Concerning flexibility, evidence is mixed. Among low-skilled workers, Mas and Pallais (2017) do not find that differences in the value for flexibility translate into gender wage gaps. In a recent randomized experiment in a large firm, Angelici and Profeta (2020) found that flexible time and space work improves the well-being and work-life balance of both male and female workers.

6. It is also important to stress that, in the absence of an experiment, estimating the willingness to pay for specific job attributes in an imperfectly competitive market has been proven difficult. Search frictions may entail small equilibrium wage differentials across jobs even in the presence of substantial preferences for amenities (Bonhomme and Jolivet, 2009). As I estimate the average firm values perceived by workers of a given group, I leave unrestricted the way wages and amenities shape worker-perceived firm values. In the model I focus on, wages and firm values are separate parameters, and I can infer the importance of non-wage components by ex-post inspecting the stationary worker-firm allocations.

2 Theoretical Framework

This section presents a theoretical framework to interpret the observed data. The objectives are twofold. First, I want to predict worker mobility across firms and into and out of non-employment. Second, I am interested in predicting log-wage distributions of workers across firms. Mobility and wage estimates allow me to assess the relative importance of key mobility components driving gender imbalances in employment across firms that translate into gender wage differentials.

I employ a finite mixture model *à la* Lentz, Piyapromdee, and Robin (2023). In what follows, I describe in detail the analysis framework and discuss the assumptions.

Agents

There are N workers and J firms. Workers are indexed by $i \in \{1, \dots, N\}$ and firms by $j \in \{0, 1, \dots, J\}$, where $j = 0$ is non-employment. Both firms and workers are heterogeneous.

Firms are associated with a finite number of K *classes*. The firm in which worker i is employed at time t is $j(i,t)$, and I denote as $k_{j(i,t)} \in \{1, \dots, K\}$ the class of firm $j(i,t)$. The class of non-employment is $k_0 = 0$.

Workers differ in their observed and unobserved characteristics. The set of observed characteristics consists of age and tenure, which are time-varying, and gender, which is time-invariant. Unobserved heterogeneity is discrete and can be clustered into L groups. Workers are thus associated with a finite number of L latent *types*, where $l_i \in \{1, \dots, L\}$ denotes workers' latent heterogeneity.

Gender interact with age and tenure to allow wages and mobility to vary between men and women over one's career within a latent type. In other words, the vector of observed characteristics x_{it} includes combinations of gender, age, and tenure observed in time period t. Each t refers to a calendar quarter.

Timing

In period 1, a worker enters the panel being employed. The initial observed heterogeneity x_{i1} determines a particular distribution of initial matches $\Pr\Big(l, k_{j(i,1)} \;\Big|\; x_{i1}\Big)$, which is left

unrestricted.

Job mobility between a firm at time t and another firm at time $t+1$ is denoted by $s_{it} = 1$. In every period $t \geq 1$, the worker changes employment status or firm class ($s_{it} = 1$ or 0) with a probability that depends on worker's type l_i, worker's characteristics x_{it}, and current firm class $k_{j(i,t)}$, denoted as $\Pr\Big(k_{j(i,t+1)} \;\Big|\; k_{j(i,t)}, l_i, x_{it}\Big)$. Transitions into and from nonemployment are left unrestricted, while job-to-job transitions are the product between a job sampling probability and a choice probability. Whether a transition occurs in the last period is unknown.

The worker draws log-wages from a static distribution that depends on worker's types and firm's classes. The distribution of log-wages is $f\Big(y_{it} \;\Big|\; l_i, x_{it}, k_{j(i,t)}\Big)$, and it is assumed to be normal with (l, x, k)-specific means and variances.

Section 2.1 formally describes all model parameters and discusses their identification.

Discussion of the assumptions

The paper aims to assess to what extent the gender wage gap is explained by men and women being sorted differently across firms, to identify the key mobility components driving gender imbalances in employment across firms, and to quantify their relative importance in determining gender wage differentials. This requires separating differences in employment opportunities from differences in job acceptance decisions. In matched employer-employee data, this separation is not feasible at the worker-firm level without additional structure.

The model describes worker mobility through three objects: offer arrival rates, acceptance decisions, and transitions into and out of non-employment. Workers compare the value of the incumbent firm with the value of the poacher, and firm-to-firm transitions occur only when an offer arrives and is accepted. Wages are drawn from match-specific distributions that are allowed to vary with worker and firm heterogeneity.

To make this framework estimable, I approximate worker and firm heterogeneity with a finite number of latent worker types and firm classes. Worker latent types interact with observed worker characteristics, allowing for a flexible relationship between observed and unobserved heterogeneity. The interpretation is that, in expectations, workers of a given type earn similar wages and have similar mobility patterns, while firms in the same class are

similar in the dimensions relevant for wages and mobility. This discretization is therefore a statistical approximation that makes the separation between opportunities and preferences feasible in the data.

Adopting a latent-type framework reduces the number of parameters to be estimated drastically, thus overcoming over-fitting issues encountered in the fixed-effect estimation proposed by Abowd, Kramarz, and Margolis (1999). The latent-type framework also improves on the fixed-effect estimation biases arising from the limited mobility of workers across firms (Bonhomme et al., 2023). Importantly, it allows us to model mobility explicitly.

Beyond its computational advantages, the latent-type structure admits theoretical interpretations. Samuelson and Steiner (2025) show that a latent representation can be viewed as an optimal decomposition of the data into coherent behavioral patterns, where each type groups agents that are best described by a common behavioural model. In this sense, latent types capture economically meaningful forms of heterogeneity rather than mere statistical clusters. This perspective aligns with work highlighting the central role of heterogeneity in shaping economic behaviour (Heckman and Singer, 1984; Bonhomme and Denis, 2025). The latent-type approach used in this paper captures systematic differences in wage and mobility responses, allowing for within and between gender heterogeneity.

Workers draw log hourly wages from a normal distribution specific to worker types, worker characteristics, and firm classes. Similar to Abowd, Kramarz, and Margolis (1999), the wage distribution does not allow for wage dynamics. Similar to Bonhomme, Lamadon, and Manresa (2019), the wage distribution does not impose additive separability between worker and firm heterogeneity. This flexibility is motivated by empirical evidence that firm wage premia need not be common across worker groups, including men and women (Card et al., 2016). Similar to Lentz, Piyapromdee, and Robin (2023), stayers and movers share the same wage means.[7]

Mobility is a Markov process independent of wage realisations conditional on worker types and firm classes. This is the standard exogenous mobility assumption (Abowd, Kramarz, and

7. In Lentz, Piyapromdee, and Robin (2023), stayers draw wages from a dynamic distribution while movers draw wages from a static distribution. The two distributions share the same mean wages but have different variances. In Abowd, Kramarz, and Margolis (1999), firm fixed effects are estimated only on movers. Bonhomme, Lamadon, and Manresa (2019) estimate wage distribution parameters (worker- and firm-specific averages and variances) only on movers.

Margolis, 1999). Exogenous mobility implies that job assignment and job-to-job mobility depend only on observed and unobserved characteristics of workers and firms. Although it rules out mobility motivated by discovering new job opportunities or, more generally, driven by idiosyncratic shocks to earnings while on the job, it still allows for different sorting patterns. In particular, I can investigate sorting patterns based on worker-firm complementarities in wages separately from sorting patterns based on preferences for non-wage components.

In the model, workers make a firm-to-firm transition if they receive a job offer and if the value of the poacher provides the worker with a higher utility than the incumbent. Firm-to-firm mobility therefore reveals preferences, while allowing for differences in the opportunity to move through the job offer arrival rate. Workers of the same type, up to an i.i.d. idiosyncratic utility draw, value firms of class $k = 1, \ldots, K$ the same.

To allow for heterogeneous reluctance to move, I introduce an inertia parameter that varies by workers' latent types and observed characteristics. This parameter shifts the relative utility of remaining with the incumbent firm and captures persistent mobility frictions that may differ systematically across worker types and across career stages. For example, workers with more experience or longer tenure may have stronger attachments to their current firm, and these attachments may also vary by ability. The idiosyncratic utility draw captures match-specific deviations in utility, while inertia captures a systematic incumbent bias common to workers in a given (l, x) group. Allowing inertia to vary across worker types avoids imposing common acceptance probabilities when workers face offers from firms in the same class, an equality restriction that can be empirically violated and distort inference by reallocating differences in mobility to other model components (Flinn and Heckman, 1982).

When choosing between the poacher and the incumbent, workers compare the common values of the two firms and an idiosyncratic utility draw. The idiosyncratic utility draw is specific to the worker-firm match and captures transitory match-specific deviations in utility, such as mobility costs that vary across potential moves. The idiosyncratic draw is distributed type I extreme value with scale parameter 1. Under this distributional assumption, upon receiving an offer from a firm in class k', acceptance increases in the ratio of the poacher's value to the incumbent's value and decreases with inertia, which scales the systematic attractiveness of remaining in the current firm. Perceived firm values and wage distributions are separate

parameters. There is no restriction on how firm values and average wages are related. This allows for the possibility that workers may value something beyond just wages in the firm.

If men and women care only about wages and earn higher wages at different firms, gender-based differences in worker-firm allocations may arise from a comparative advantage explanation. If women instead care about amenities more than wages, then they may be more likely to sort into firms that offer higher levels of amenities, which may not necessarily be the ones that would pay them more.

The job ladder based on utility levels closely mirrors the one proposed by Sorkin (2018), who analyses firm-to-firm transitions to estimate utility levels of working at a firm and compares them to firm-level earnings to find the role played by compensating differentials in explaining wage inequality. Sorkin (2017) adopts the same revealed preference estimation technique to study the gender wage gap in US. The adoption of the finite mixture approach permits heterogeneous workers both within gender and across gender and to be more in line with key features of theoretical sorting models. For example, Bagger and Lentz (2019) view job-to-job moves as a revelation of preferences in a framework that allows for worker heterogeneity in skill levels, while Taber and Vejlin (2020) also use a revealed preference argument, highlighting the importance of preferences for non-wage components in determining worker choices between two jobs.

To sum up, from a theoretical point of view, the latent-type model relates to partial equilibrium on-the-job search models with heterogeneous workers and firms, wage posting, random preferences for job types, and worker-specific offer arrival rates. From an empirical point of view, if I impose additivity between worker and firm heterogeneity in the wage equation, the model reduces to a latent-type version of Abowd, Kramarz, and Margolis (1999).

2.1 The theoretical framework in practice

The observed data for worker i consist of sequences of firm identifiers $\Big(j(i,1),\ldots,j(i,T)\Big)$, log-hourly wages $\Big(y_{i1},\ldots,y_{iT}\Big)$, mobility indicators $\Big(s_{i,1},\ldots,s_{i,T-1}\Big)$, gender, and time-varying tenure and age categories. Interactions between gender and tenure and age categories are collected in a time-varying vector x_{it}. The latent data consist of the unobserved heterogeneity

types $l_i \in \{1, \ldots, L\}$ and $k_{j(i,t)} \in \{0, 1, \ldots, K\}$, for $i \in \{1, \ldots, N\}$, $j \in \{0, 1, \ldots, J\}$, and $t \in \{1, \ldots, T\}$.

Conditional on a classification C of firms into classes, on the initial characteristics x_{i1}, and on a value θ of the parameters, the complete likelihood of worker i's history is as follows.

$$\begin{aligned}\mathcal{L}_i(\theta|l_i, x_{i1}, C) = \Pr\Big(l_i, k_{j(i,1)} \;\Big|\; x_{i1}\Big) \times \prod_{t=1}^{T-1} \Bigg\{ \Pr\Big(k_{j(i,t+1)} \;\Big|\; k_{j(i,t)}, l_i, x_{it}\Big)^{\mathbb{1}\{s_{it}=1\}} \\ \times \Pr\Big(\neg \;\Big|\; k_{j(i,t)}, l_i, x_{it}\Big)^{\mathbb{1}\{s_{it}=0\}} \Bigg\} \\ \times \prod_{t=1}^{T} f\Big(y_{it} \;\Big|\; l_i, x_{it}, k_{j(i,t)}\Big)\end{aligned} \tag{1}$$

The likelihood function factors into three parts: contributions from the initial matching distribution, contributions from the mobility processes, and contributions from hourly wages.

Initial matching distribution

At $t = 1$, worker i enters the panel being employed. Observed characteristics, x_{i1}, determine the initial probability of worker-firm match $\Pr\Big(l_i, k_{j(i,1)} \;\Big|\; x_{i1}\Big)$. The worker's observed characteristics consist of interactions between her gender, $g_i \in \{F, M\}$, and combinations of short/long tenure status and age groups. Following Lentz, Piyapromdee, and Robin (2023), short tenure is defined to be less than two years in employment and less than six months in non-employment. Age is divided into five categories: 25-30, 31-35, 36-40, 41-45, and 45+. The resulting vector of observed characteristics includes $2 \times 2 \times 5$ categories. The initial matching parameter is left completely unrestricted and estimated using simple frequencies. For notational simplicity, I denote the initial matching distribution as $m_0\Big(l, k \;\Big|\; x\Big)$.

Within a latent class, each firm is equally likely to be selected. I abstract from modeling firm-specific sampling weights in the likelihood. Conditional on a firm's class assignment, this component simplifies out in the expectation step of the EM algorithm used to compute the posterior probability that worker i is of type l. In the maximisation step, it enters the log-likelihood additively as a separable parameter and therefore does not affect the estimation of the type-specific parameters of interest. The uniform sampling assumption does

not affect identification and estimation identification and estimation of the latent structure. Moreover, because the estimation proceeds in two stages, first clustering firms and then clustering workers conditional on the firm classification, in principle any assumption about the firm-specific sampling can be made. Section 3.2 further clarifies this point.

Mobility processes

At each period $t \in \{1, \dots, T-1\}$, I observe whether the worker separates from the current firm, $s_{it} = 1$, or stays, $s_{it} = 0$. Mobility at $t = T$ is unknown. The worker changes employment status and firm class with a probability that depends on worker's type, worker's characteristics, and current firm class, $\Pr\Big(k_{j(i,t+1)} \;\Big|\; k_{j(i,t)}, l_i, x_{it}\Big)$. For notational simplicity, denote the current firm class by k and subsequent firm class by k'. In addition, denote the transition probability by $m\Big(k' \;\Big|\; k, l, x\Big)$. The worker stays with probability $m\Big(\neg \;\Big|\; k, l, x\Big) = 1 - \sum_{k'=0}^{K} m\Big(k' \;\Big|\; k, l, x\Big)$.

Job-to-Job transitions. At any time period t, a worker of type l and with characteristics x_{it} moves from a firm of class $k = 1, \dots, K$ to a firm of class $k' = 1, \dots, K$ if the worker receives an offer and if prefers the poacher over the incumbent. The poacher is preferred if the perceived value of the match (l, x, k') is higher than the perceived value of the match (l, x, k). The probability of a job-to-job transition is specified as the product of an offer arrival probability and a choice probability,

$$m\Big(k' \;\Big|\; k, l, x\Big) = \lambda_{lxk'}\; P_{lx}\Big(k' \succ k\Big)$$

where $\lambda_{lxk'}$ represents the probability that a worker of type (l, x) receives an offer by a different firm of class k'.[8]

Upon receiving an offer, the worker compares the common values of the poacher and the incumbent together with an idiosyncratic utility draw. The worker moves if the firm of class k' is preferred over the firm of class k. I introduce a worker-type-specific inertia parameter $h_{lx} > 0$ that scales the effective value of remaining with the incumbent. Assuming the idiosyncratic draw is distributed according to a type I extreme value, the probability of accepting an offer from k' while employed in class k is

8. Note that the different firm may belong to the same class of the firm in the current period.

$$P_{lx}\Big(k' \succ k\Big) = \frac{\gamma_{lxk'}}{\gamma_{lxk'} + h_{lx}\,\gamma_{lxk}} = \frac{\gamma_{lxk'}/\gamma_{lxk}}{\gamma_{lxk'}/\gamma_{lxk} + h_{lx}}$$

Acceptance therefore increases in the ratio $\gamma_{lxk'}/\gamma_{lxk}$ and decreases with inertia h_{lx}. The parameter h_{lx} acts as a mobility wedge that operates on the accept/reject margin. To be precise, γ_{lxk} is a monotonic transformation of the firm values, and $h_{lx} = \exp(\alpha_{lx})$ is a monotonic transformation of the inertia parameter.

For the estimation, the choice probability retains a Bradley-Terry structure (Bradley and Terry, 1952; Hunter, 2004), modified to account for inertia (Davidson and Beaver, 1977). The Bradley-Terry specification was initially introduced to model a situation in which individuals are repeatedly compared with one another in pairs. As matched employer-employee data can be represented in a directed graph where nodes are firms and edges are non-negative integers of worker transitions between any pair of firms, the Bradley-Terry specification is useful to estimate how workers value firms under the assumption that they make only binary choices.

Using information on relative flows between firms, it is possible to recover a worker-specific ranking of firm classes. For each (l, x) group, I build a directed $K \times K$ transition matrix whose (k, k') cell records transitions from class k to class k'. Identification relies on the usual connectivity condition: for each pair of classes, there must be a path linking them through observed mobility flows. Under this condition, the vector of perceived firm values $\gamma_{lx} = (\gamma_{lx1}, \ldots, \gamma_{lxK})$ is identified up to scale. Two perceived value vectors γ_{lx} and γ'_{lx} are equivalent if one is a scalar multiple of the other. The firm values are thus normalised so that $\sum_{k=1}^{K} \gamma_{lxk} = 1$. What matters for mobility are relative values across classes, not levels.

Identification proceeds in three steps under three conditions: (i) offer rates are destination-specific, (ii) the mobility graph over classes for each (l, x) is connected, and (iii) there is enough bilateral asymmetry in cross-class flows. In practice, separation of γ_{lx} and h_{lx} is not informative with only two classes, and can remain weakly informative when transition matrices are sparse or nearly symmetric.

First, within-class transitions satisfy

$$m(k \mid k, l, x) = \frac{\lambda_{lxk}}{1 + h_{lx}}.$$

Hence, at this stage we identify $\lambda_{lxk}/(1+h_{lx})$, and therefore offer-rate ratios,

$$\frac{\lambda_{lxk}}{\lambda_{lx\tilde{k}}} = \frac{m(k \mid k, l, x)}{m(\tilde{k} \mid \tilde{k}, l, x)}.$$

Second, cross-class directional asymmetries satisfy

$$\frac{m(k' \mid k, l, x)}{m(k \mid k', l, x)} = \frac{\lambda_{lxk'}}{\lambda_{lxk}} \cdot \frac{\gamma_{lxk'}}{\gamma_{lxk}} \cdot \frac{\gamma_{lxk} + h_{lx}\gamma_{lxk'}}{\gamma_{lxk'} + h_{lx}\gamma_{lxk}}.$$

After substituting the identified λ-ratios, these moments and the normalization $\sum_{k=1}^{K} \gamma_{lxk} = 1$ define an over-identified non-linear system in (γ_{lx}, h_{lx}). Intuitively, h_{lx} is a common scalar wedge across all bilateral comparisons, while γ_{lx} is class-specific. Cross-equation consistency across multiple class pairs is what makes h_{lx} informative.

To see how h_{lx} enters, define

$$A_{kk'} \equiv \frac{m(k' \mid k, l, x)}{m(k \mid k', l, x)} \cdot \frac{\lambda_{lxk}}{\lambda_{lxk'}},$$

which is observed once λ-ratios are recovered from within-class transitions. Let $r_{kk'} \equiv \frac{\gamma_{lxk'}}{\gamma_{lxk}}$. Then

$$A_{kk'} = \frac{r_{kk'}(1 + h_{lx} r_{kk'})}{r_{kk'} + h_{lx}}$$

implies the quadratic equation

$$h_{lx} r_{kk'}^2 + (1 - A_{kk'}) r_{kk'} - A_{kk'} h_{lx} = 0.$$

Therefore, for any fixed $h_{lx} > 0$, each pair (k, k') delivers a unique positive $r_{kk'}$. Values of h_{lx} that are consistent with the data are those for which these pairwise ratios can be embedded in a common vector γ_{lx} across classes, for example by satisfying $r_{ab} r_{bc} = r_{ac}$ for all triples (a, b, c). In this sense, the data identify λ-ratios and place over-identifying restrictions on the admissible pairs (γ_{lx}, h_{lx}). Once a value of h_{lx} is selected, γ_{lx} is recovered from the $r_{kk'}$'s and the normalization $\sum_k \gamma_{lxk} = 1$. Offer-rate levels then follow from within-class transitions,

$$\lambda_{lxk} = (1 + h_{lx})\, m(k \mid k, l, x).$$

Appendix D details the MM routine: γ is updated with iterative Bradley-Terry steps conditional on h, while $h_{lx} = \exp(\alpha_{lx})$ is updated by solving a one-dimensional first-order condition and selecting the unique positive root via bracketing and bisection. The system is over-identified, but identification of h_{lx} is weak when transition matrices are sparse or nearly symmetric, in which case many (γ, h) pairs can be observationally close in finite samples.[9]

The identification strategy also relates to recent work by Arcidiacono et al. (2025). They study a dynamic job search model with preference shocks and show how observed transition hazards can be used to separately recover offer arrival rates and job acceptance probabilities. Their key insight is that preference shocks create a link between value functions and conditional choice probabilities, so that observed mobility hazards reveal differences in the values workers assign to jobs.

The approach of this paper relies directly on observed bilateral mobility flows within worker types. Conditional on meeting, if a higher number of workers of type l with characteristics x move from k to k' than from k' to k, then, after accounting for offer arrival rates, we infer that this group values firms of class k' more than firms of class k. This revealed-preference logic, implemented in a finite-mixture framework, underlies the estimation of worker-specific firm values and defines what is meant by preferences throughout the paper.

Transitions to and from non-employment. At any time period $t \in \{1, \dots, T-1\}$, a worker of type l and with characteristics x moves from a firm of class $k = 1, \dots, K$ to non-employment $k = 0$ with probability $m\Big(0 \;\Big|\; k, l, x\Big) = \delta_{lxk}$. The worker moves from non-employment to a firm of class $k' = 1, \dots, K$ with probability $m\Big(k' \;\Big|\; 0, l, x\Big) = \psi_{lxk'}$.

Transitions to and from non-employment are left completely unrestricted, and are identified by simple frequencies. Moving into non-employment depends on worker type and on current firm class, moving into employment depends on worker type and new firm class. $m\Big(0 \;\Big|\; 0, l, x\Big) = 0$ as there are no transitions from non-employment into non-employment. Given the specification of the transition probability parameters, the probability of staying

9. To stabilize estimation in small cells, I impose a small lower bound on each γ_{lxk} at each MM iteration and then renormalize. This avoids degenerate zero-value classes that create numerical instability in the α and transition updates, while leaving identification driven by relative flow asymmetries.

into non-employment is as follows.

$$m\Big(\neg \mid 0, l, x\Big) = 1 - \sum_{k'=1}^{K} \psi_{lxk'}$$

For employed workers, $k \geq 1$, the probability of staying with the same firm is as follows.

$$m\Big(\neg \mid k, l, x\Big) = 1 - \delta_{lxk} - \sum_{k'=1}^{K} \left(\lambda_{lxk'} \, \frac{\gamma_{lxk'}}{\gamma_{lxk'} + h_{lx}\gamma_{lxk}} \right)$$

Hourly wage distributions

Hourly wages are drawn from a static worker-firm-specific log-normal distribution.

$$\log f\Big(y_{it} \;\Big|\; l, x, k\Big) = -\log(\sigma_{lxk}) - \log(\sqrt{2\pi}) - \frac{1}{2}\left(\frac{y_{it} - \mu_{lxk}}{\sigma_{lxk}}\right)^2$$

Earnings and hours are recorded at annual frequency. There is only one payroll recorded for each employment spell in a year. I calculate hourly wages by dividing annual earnings by the number of hours and consider that the hourly wage for a given quarter. Estimates of μ_{lxk} and σ_{lxk}, for any match (l, x, k), will be used to compute the gender wage gap in a framework that allows for earnings complementarities between workers and firms.

3 Estimation

Firm classes and worker types are unobserved. The mobility of workers across firm types makes it difficult to separate the complete log-likelihood across firm types. Therefore, I proceed with a two-step estimation as in Bonhomme, Lamadon, and Manresa (2019). First, I cluster firms into classes using a k-means algorithm. Second, conditional on the firm clustering, I use an Expectation-Maximisation (EM) algorithm to iterate over (a) the calculation of the posterior probability that worker i is of type $l = 1, \ldots, L$, and (b) the maximisation of the expected log-likelihood with respect to the parameters of interest.[10]

10. In all fairness, Lentz, Piyapromdee, and Robin (2023) classify both firms and workers in the EM algorithm. They treat firm types as parameters to be estimated in the expected likelihood maximisation along with the other parameters. This has the advantage of fully using the information on both wages and mobility for both workers and firms.

3.1 K-Means Algorithm to Cluster Firms

In the model described in section 2, the initial matching distributions, log wages, and mobility patterns depend on firm classes but not directly on firm identities. The idea is that unobserved firm heterogeneity is captured at the class level and not at the individual firm level (Bonhomme, Lamadon, and Manresa, 2019). Therefore, I partition the J firms into a finite number of classes, K, solving a weighted k-means problem: I use as input characteristics of each firm male and female empirical cumulative distribution functions of log-hourly wages and female shares, and I weight by average firm size. As I want to estimate earnings distributions of male and female workers and their mobility patterns across firm types, I need the k-means algorithm to take care of firms' behaviour towards a specific gender.

I residualise log-hourly wages on 3-digit occupational and year dummies. These effects are estimated on the female sample only, and both male and female log-hourly wages are purged using these effects. I do so to control for observed workers' skills, proxied by occupations, without imposing similar (or different) returns between men and women. This is a way to remove wage differences solely due to occupational differences between men and women (Blau, Brummund, and Liu, 2012). The EM algorithm is performed on log-hourly wages residualised on the same fixed effects estimates.

I choose $K = 10$, seeking a balance between minimising total intra-class variation and ensuring sufficient observations to fit the data well. Appendix C presents more details about the k-means algorithm and cluster-stability diagnostics.

3.2 EM Algorithm to Classify Workers

Worker types are unobserved. The EM algorithm classifies workers into a discrete number of types L by iterating an expectation step and a maximisation step until convergence. I set $L = 3$ and show the fit of the model in Appendix Figure B1.[11]

11. Although the small number of points of support is computationally convenient, it can be shown that just a few points of support approximate well the underlying distribution of unit fixed effects and their correlation with covariates. I thank Seth Sanders for the stimulating discussion. In addition, simulations show that precision deteriorates sharply when some cells are sparsely populated. Figure B1 shows that the selected number of types fits the data well.

Expectation step

For given parameters $\theta^{(m)}$ and a firm classification C, I compute the posterior probability that worker i is of type $l = 1, \ldots, L$.[12]

$$p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) = \frac{\mathcal{L}_i\Big(\theta^{(m)} \mid l_i = l, x_{i1}, C\Big)}{\sum_{l'=1}^{L} \mathcal{L}_i(\theta^{(m)} \mid l_i = l', x_{i1}, C)}. \tag{2}$$

Maximisation step

The M-step maximises the expected complete-data log-likelihood with respect to θ:

$$\sum_{i=1}^{N} \sum_{l=1}^{L} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \log \mathcal{L}_i(\theta \mid l_i = l, x_{i1}, C)\,.$$

This yields the updated parameter vector $\theta^{(m+1)}$, which is then used in the next E-step. I iterate between the E-step and the M-step until convergence. For the wage distributions, the M-step updates are weighted means and variances, with posterior probabilities as weights. For the initial matching distribution and the unrestricted transition probabilities, the updates are weighted frequencies. Transition probabilities for job-to-job mobility are non-linear in the parameters of interest. Appendix D details the minorisation-maximisation (MM) algorithm proposed by Lentz, Piyapromdee, and Robin (2023) for these non-linear components and explains how I adapt it to estimate inertia. In particular, it details the updates for γ, λ, and δ, and shows how inertia $h_{lx} = \exp(\alpha_{lx})$ is estimated within the EM iterations.

4 Data and Sample Selection

I use French matched employer-employee data from the *Déclarations Annuelles de Données Sociales* (DADS) over 2015-2019. These datasets are collections of mandatory employer payroll declarations collected by the French statistical institute, *Institut National de la Statistique et des Études Économiques* (INSEE), and provide job-spell-level information on worker-firm matches. Importantly, working hours are reported, allowing me to account for

12. Any individual firm sampling factor in equation 1, conditional on a firm classification C, would cancel out in equation 2.

gender differences in labour supply. The core analysis relies on DADS-Postes, complemented with the *Conditions de Travail* surveys (CT-RPS, 2016 and 2019), which contain information on working time, schedule demands, and workplace conditions.[13]

Sample selection. DADS-Postes covers the universe of jobs in France. In its raw form, it lacks a proper longitudinal dimension because worker identifiers change every two years. I therefore recover worker histories by linking annual files using the procedure in Palladino et al. (2025a). For each job spell, I have information on worker identifier, firm identifier, year, yearly earnings, hours worked, occupation, worker age, worker tenure in the firm, and the starting and ending days of the spell in the year. I consider any job with positive wages and hours as employment. Wages are reported at the annual spell level. I therefore construct an hourly wage by dividing yearly earnings by yearly hours worked and assign that hourly wage to the quarters covered by the spell. Mobility is tracked at quarterly frequency using changes in employment status and firm identifiers across spells.

I first define the firm sample by selecting firms that employ at least one full-time worker in *Ile-de-France* in 2015, employ both genders, and are active in each year from 2015 to 2019. I then construct worker trajectories and retain workers employed in the first quarter of 2015, only working in *Ile-de-France* over the period 2015-2019, never working in the Agricultural sector, holding part-time and full-time contracts that last at least a month, who are between 25 and 55 years old, and who have worked only in the selected firms. I track the selected workers over time. Non-employment is the time between different job spells.[14]

Appendix Table A1 compares firms across different selection steps. Moving from the broad 2015 universe (Sample 1) to firms that employ both genders and remain active for five years (Sample 2) substantially reduces the number of firms and shifts the sample toward larger, higher-paying employers. Column 3 reports the firms where I observe workers selected for the analysis. Compared with Sample 1, these firms are larger (83 vs 21 employees on average), pay higher annual earnings (EUR 27,799 vs EUR 14,287), have longer working

13. CT-RPS is a survey. When linked at the firm level to DADS-Postes, the matched sample covers about 2% of firms; these data are used for supplementary descriptive evidence on workplace conditions.

14. Non-employment does not include periods of maternity leave in a firm or retirement, but could encompass periods of inactivity. As I do not observe education, I consider only workers aged 25+ to minimize the risk that non-employment is confounded with periods of education.

hours (1,329 vs 793), and display a more balanced gender composition.

Appendix Table A2 reports worker characteristics across selection steps. The restrictions mainly reduce sample size and leave composition broadly stable: average age remains about 40, and tenure increases only modestly (from 6 to 7 years). Relative to the initial sample, the final sample has slightly higher earnings and hours for both genders (women: EUR 31,213 to 32,565 and 1,417 to 1,552 hours; men: EUR 38,676 to 39,940 and 1,483 to 1,644 hours). The final sample includes 1,066,931 women and 1,085,686 men.

I focus on *Ile-de-France* to reduce the sample size. As type and class effects are time-invariant, I focus on a five-year short panel. After all, assuming that a worker's unobserved ability is constant over long periods is deemed unrealistic. The same reasoning applies to job arrival rates.

Sample description. Table 1 presents summary statistics for the estimation sample, which is balanced by gender. Relative to men, women earn EUR 7,834 less on average (22.4% of the female mean) and work 95 fewer hours per year, consistent with a higher prevalence of part-time employment. Monthly mobility rates are otherwise similar across genders, although women are more likely to transition from non-employment to employment. Occupational composition also differs: women are more likely to be intermediate managers and non-manual employees, and the opposite is true for managerial and manual employment positions. The average job spell duration is similar across both groups, at roughly 3.5 years. Appendix Table A3 reports the corresponding breakdown by gender, age, and tenure.

Figure 1 provides descriptive evidence on gender differences in earnings, mobility patterns, and destination pay over the life cycle. Panel (a) shows that earnings increase with age, although more so for men than women. Among workers aged 25-30, women earn EUR 3,000 less than men, a gap that widens to EUR 10,000 among workers aged 55+.

Panel (b) plots the distribution of destination-firm average pay for hires from non-employment and shows virtually perfect overlap between men and women. Panel (c) shows that, among job-to-job transitions, men are more likely than women to make an upward move, defined as a transition to a firm with a higher firm-level mean log hourly wage. This gap closes by age 47. Panel (d) reports the same statistic for transitions from non-employment

to employment: the male advantage is large at younger ages, shrinks over the life cycle, and, if anything, reverses around age 45. Overall, the figure documents the well-known widening gender earnings profile over the life cycle (Goldin, 2014; Goldin, Kerr, Olivetti, and Barth, 2017), together with clear gender differences in upward mobility early in the career. Even though destination-pay distributions at hiring from non-employment are similar across genders, men's higher upward mobility in job-to-job and early re-entry transitions can cumulate into larger earnings gaps and diverging career trajectories over time.

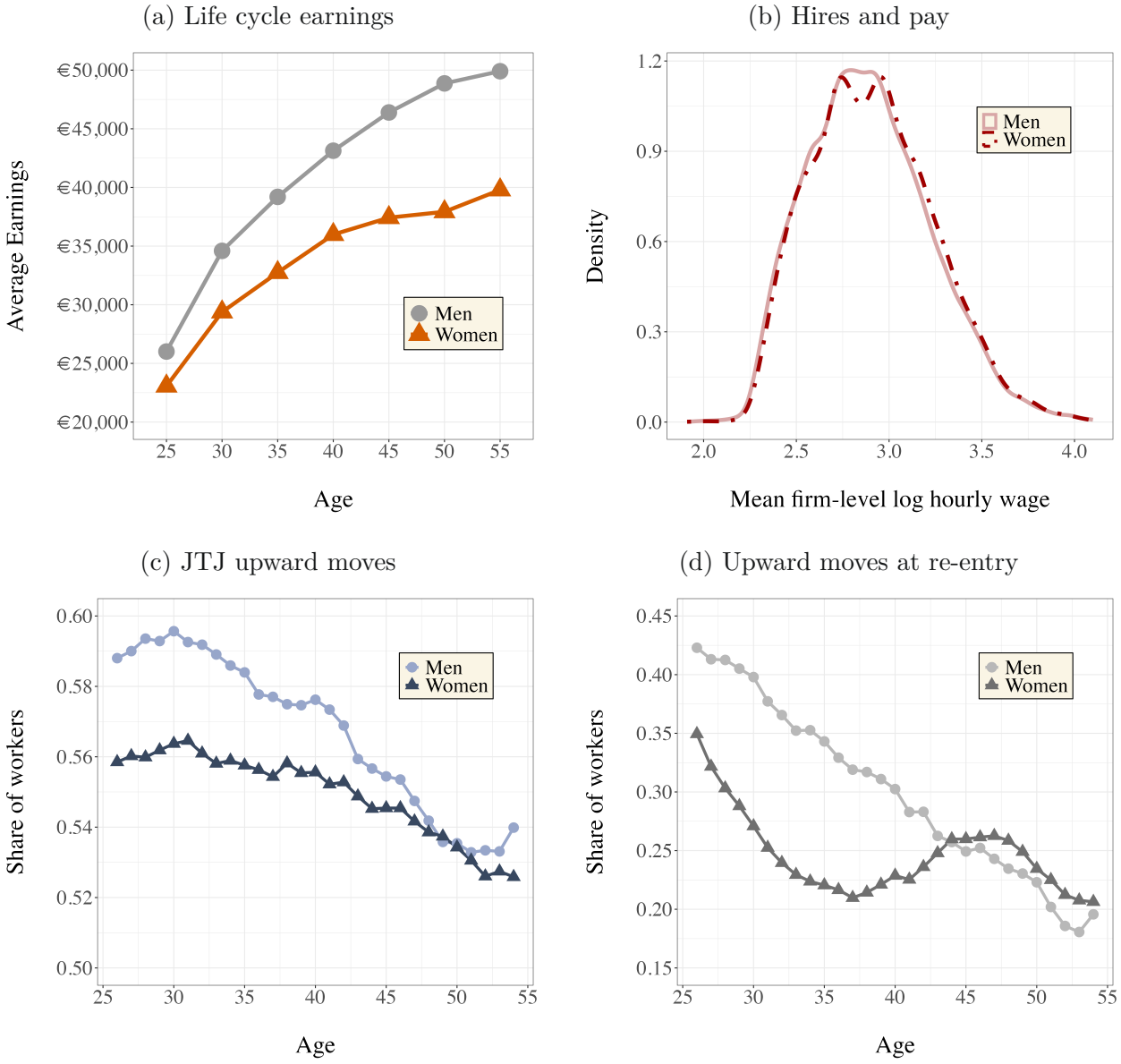
Taken together, these patterns suggest that the relevant gender differences may lie less in whether workers move and more in the types of firms they move to, especially early in the career. The next section explores this hypothesis by documenting how the estimated model maps workers into latent groups with distinct wage profiles, mobility patterns, and preferences.

Table 1: Sample description

Gender:	Women	Men
Mean Annual Earnings	34,915	42,749
Mean Hours	1,661	1,756
Share Part-time	17%	4%
Mean Age	40	41
Share doing JTJ	15%	14%
Share doing E-NE	34%	34%
Share doing NE-E	17%	11%
Share Managers	24%	28%
Share Intermediate	30%	22%
Share Employee non-manual	38%	21%
Share Employee manual	8%	28%
Mean job spell quarters	13	13
N workers	1,066,802	1,085,505

Notes: This table reports descriptive statistics for the estimation sample over 2015Q1–2019Q3. Mobility measures are the share of workers who experience at least one transition of the indicated type during the sample period. *JTJ* denotes job-to-job transitions; *E* and *NE* denote employment and non-employment, respectively. All remaining statistics are sample averages computed over the pooled period.

Figure 1: Gender differences in life cycle earnings, upward mobility, and destination pay



Notes: The figure compares gender differences in life cycle earnings, destination pay, and upward mobility. Panel (a) reports age-specific unconditional average annual earnings (EUR) for men and women in the analysis sample. Panel (b) reports kernel densities of the firm-level mean log hourly wage at destination firms for hires from non-employment, by gender. Panel (c) reports, by age and gender, the share of job-to-job transitions that are upward, where an upward move is defined as moving to a firm with a higher firm-level mean log hourly wage than the origin firm. Panel (d) reports the analogous share for transitions from non-employment to employment, with upward re-entry defined relative to the worker's last observed employer before non-employment. Firm-level means are computed from the pooled 2015–2019 sample. Age mobility profiles are shown as centered 3-year rolling averages.

5 Firm Classes and Worker Types

This section shows that the estimated latent firm classes and worker types capture economically meaningful heterogeneity in wages, mobility, and preferences. I begin with ex post tabulations of observed firm and worker characteristics by predicted latent class and type, and then use the estimated preference parameters to characterize heterogeneity within and across genders. To facilitate interpretation, I relabel firm classes and worker types so that they are ordered by increasing estimated log hourly wages $\hat{\mu}_{lkx}$.[15] Throughout this section, higher latent groups are therefore interpreted as higher-paying firms and higher-wage workers.

Firm classes. Table 2 reports descriptive statistics for the average firm in each predicted class over the pooled 2015–2019 sample. The classes display substantial heterogeneity in both size and workforce composition. Mean firm size ranges from 28 to 156 workers and is largest in the middle classes, while gender composition varies sharply across classes, with female shares ranging from 22% to 70%.

Despite this heterogeneity, gender composition does not cleanly map onto wages, hours, or industry sectors. By construction, higher classes are higher-paying, and the pay gradient is increasing for both genders, implying that firms that pay more do so for both groups. Higher-paying classes tend to display longer annual hours, and sectors such as construction and finance are relatively more represented at the top. In contrast, hospitality, education, and health are more concentrated in lower-paying classes.

Managerial composition also varies sharply across classes, but again not mechanically with female workforce shares. In some classes, women are more likely than men to be managers, whereas in others, the opposite is true. Even in female-majority classes, women are often underrepresented among managers relative to their workforce share, whereas in some male-dominated classes they are proportionally represented or slightly overrepresented.

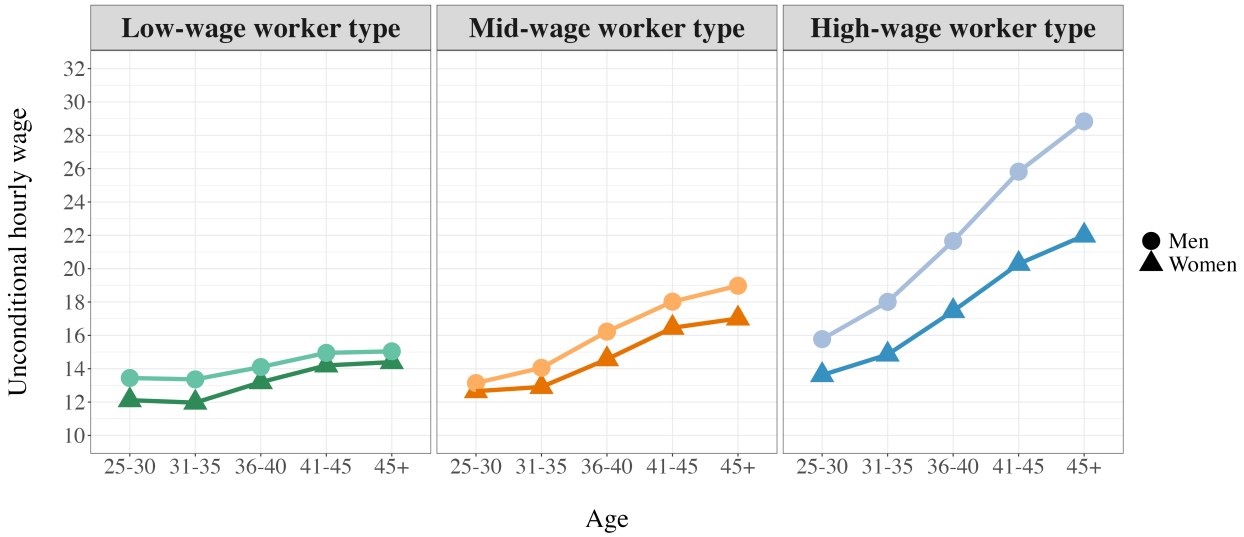
15. Latent worker types l and latent firm classes k are, per se, meaningless; they capture unobserved heterogeneity of groups of workers and firms. To relabel them, I run the standard two-way projection $Y = \nu_x + \alpha_l + \psi_k + \epsilon$ on a cross-section drawn from the empirical matching distribution $\Pr(l, x, k)$, where ν_x includes interactions of observable characteristics, α_l is the worker-type effect and ψ_k is the firm-class effect. The matching distribution combines the stationary worker-firm allocation $\Pr^*(k \mid l, x)$ with the empirical frequency of worker groups (l, x). I then reorder l and k so that α_l and ψ_k are increasing. The relabelling allows interpreting higher l as higher-wage worker types and higher k as higher-paying firm classes.

Table 2: Firm classes description

Firm class:	1	2	3	4	5	6	7	8	9	10
N firms	19,246	15,090	15,443	8,487	9,377	7,499	12,065	13,037	11,714	14,629
Mean size	28	57	64	152	156	143	94	102	85	61
Female share	39%	70%	27%	66%	26%	62%	61%	22%	68%	40%
Mean log hourly wage (EUR)	2.49	2.59	2.68	2.63	2.73	2.82	2.94	2.95	3.06	3.28
Mean hours	1,200	1,192	1,270	1,212	1,301	1,339	1,293	1,365	1,307	1,398
Mean log hourly wage (EUR) - women	2.45	2.64	2.61	2.62	2.72	2.80	2.86	2.92	2.98	3.22
Mean log hourly wage (EUR) - men	2.53	2.55	2.75	2.63	2.74	2.84	3.03	3.00	3.21	3.38
Share of women managers	8%	13%	6%	6%	4%	9%	20%	6%	18%	12%
Share of men managers	13%	9%	14%	4%	11%	8%	18%	18%	14%	20%
Share of women among managers	35%	59%	29%	57%	25%	53%	52%	25%	56%	36%
Share of firms in Hotel	10%	14%	16%	27%	24%	16%	2%	8%	3%	3%
Share of firms in Admin Services	5%	6%	6%	9%	9%	9%	5%	6%	8%	8%
Share of firms in Construction	8%	0%	16%	0%	9%	1%	2%	25%	2%	17%
Share of firms in Commerce	26%	20%	20%	18%	20%	20%	18%	20%	22%	22%
Share of firms in Education	3%	9%	1%	3%	1%	2%	8%	1%	4%	1%
Share of firms in Managing	11%	13%	8%	5%	5%	10%	19%	8%	17%	12%
Share of firms in Finance	2%	3%	1%	2%	1%	4%	7%	2%	8%	9%
Share of firms in Pub Admin	0%	2%	0%	8%	0%	4%	1%	0%	2%	1%
Share of firms in Health Accomm	1%	4%	0%	7%	0%	6%	2%	0%	2%	0%

Notes: This table reports ex-post tabulations of observed firm characteristics by predicted latent firm class. Firm classes are obtained from the k-means procedure described in Section 3.1 and are relabelled so that higher class indices correspond to higher-paying classes. All statistics are computed over the full sample period, 2015–2019.

Figure 2: Hourly wages by worker type



Notes: The figure reports unconditional average hourly wages by age, gender, and predicted worker type. The three panels correspond to low-, mid-, and high-wage worker types. Age is grouped into five categories (25–30, 31–35, 36–40, 41–45, and 45+). Workers are classified using the EM algorithm described in Section 3.2. Statistics are computed over the pooled 2015–2019 sample and weighted by the estimated posterior type probabilities, $p_i\left(l\middle|\theta^{(m)}, x_{i1}, C\right)$.

Worker types. Figure 2 reports unconditional hourly wages by age, gender, and latent worker type, and shows that latent types capture distinct career paths. Workers in the low-wage type exhibit nearly flat wage profiles over the life cycle, with only modest increases from ages 25–30 to 45+ (women: EUR 12.1 to 14.4; men: EUR 13.4 to 15.0). By contrast, workers in the mid-wage and high-wage types display clear upward wage trajectories, especially in the high-wage type. For the mid-wage type, wages rise from EUR 12.7 to 17.0 for women and from EUR 13.1 to 19.0 for men; for the high-wage type, they rise from EUR 13.6 to 22.0 for women and from EUR 15.8 to 28.8 for men. These patterns imply that the gender wage gap narrows over the life cycle in the low-wage type, but widens in the mid- and, especially, high-wage types.

Appendix Table A4 shows that, within types, mobility declines with age and is broadly comparable across genders. Higher-wage types are more mobile across firms, especially at younger ages, and less likely to exit into non-employment. Appendix Tables A5 and A6 report the female representation and the occupational composition breakdown, respectively.

Table 3: Heterogeneity in worker preferences

	Panel A: Within-gender rank deviation in preferences γ									
Tenure:	Short					Long				
Age:	25-30	31-35	36-40	41-45	45+	25-30	31-35	36-40	41-45	45+
Men	0.73	0.80	0.73	0.80	0.73	1.60	1.47	1.00	0.87	0.80
Women	1.53	1.07	1.47	1.20	1.00	1.20	1.33	1.33	0.80	0.80
	Panel B: Cross-gender rank deviation in preferences γ									
Tenure:	Short					Long				
Age:	25-30	31-35	36-40	41-45	45+	25-30	31-35	36-40	41-45	45+
Women vs Men	1.87	1.80	2.07	1.87	1.67	2.73	2.67	1.60	1.20	0.73

Notes: The table reports heterogeneity in estimated preferences over firm classes. Within each age-tenure cell, firm classes are ranked by estimated preference parameters for each worker type, separately for women and men. Panel A measures within-gender disagreement across worker types, using the average distance between each type-specific ranking and the gender-specific Borda consensus ranking. Panel B measures cross-gender disagreement by comparing women's and men's rankings within worker type, then averaging across types. Higher values indicate greater disagreement in preference orderings across firm classes.

Preferences. Do worker types differ in their preferences over firm classes? Table 3 reports two rank-based measures of disagreement in γ-implied firm rankings. Panel A measures within-gender heterogeneity across worker types. For each gender $g \in \{F, M\}$ and age-tenure cell x, firm classes k are ranked within each worker type l, yielding r_{glxk}. A Borda consensus is then constructed within (g, x) by $S_{gxk} = \sum_l r_{glxk}$, with consensus rank r^B_{gxk} obtained by re-ranking S_{gxk}, where lower S_{gxk} implies a better consensus rank. The statistic is

$$D^{\text{within}}_{gx} = \frac{1}{L}\sum_{l=1}^{L}\frac{1}{K}\sum_{k=1}^{K}\left|r_{glxk} - r^B_{gxk}\right|.$$

Higher values indicate greater disagreement across latent worker types within gender. The table shows that latent types capture distinct preference orderings over firms, not only different wage profiles. Panel B measures cross-gender disagreement within worker type:

$$D^{\text{cross}}_{x} = \frac{1}{L}\sum_{l=1}^{L}\frac{1}{K}\sum_{k=1}^{K}\left|r^{\text{Fem}}_{lxk} - r^{\text{Men}}_{lxk}\right|.$$

This isolates how differently men and women rank the same firm classes, holding (l, x) fixed. Cross-gender divergence is highest early in the career and declines with age. The magnitudes are sizeable: within-gender disagreement often ranges from about 0.7 to 1.6 rank positions, while cross-gender disagreement ranges from 0.7 to 2.7. Since the maximum possible value of D_x^{cross} with 10 firm classes is 5, the largest values correspond to more than half of the theoretical maximum.

Table 4 provides suggestive evidence on what these preference differences may reflect. Focusing on correlations between estimated preferences and non-wage firm attributes, conditional on wages, the clearest pattern is a sharp gender divergence among high-wage workers at early career stages. At ages 25–30 and 31–35, correlations between preferences and working hours are -0.29 and -0.12 for women, compared with 0.47 and 0.38 for men; the corresponding correlations with overtime incidence are -0.06 and 0.08 for women, versus 0.44 and 0.41 for men. These correlations are descriptive and based on the CT-RPS-linked subsample, which covers about 2% of firms per class, so they should be interpreted with caution. Still, because the CT-RPS information is aggregated to the firm-class level, these correlations remain informative about broad directional differences in how women and men value observable job attributes across classes. Importantly, these patterns are estimated conditional on wages, so they should not be interpreted as indicating that women have weaker preferences for pay. Rather, they indicate different valuations of non-wage firm attributes among firms with similar wage components.

Appendix Table A7 shows that, across firm classes, both working hours and overtime incidence are positively correlated with wages in nearly all worker-type and age cells, with correlations typically in the 0.3–0.7 range and often larger in mid- and high-wage groups. Thus, the attributes for which early-career high-wage women display weaker preferences are also those most strongly associated with higher-paying firm classes.

Taken together, these patterns suggest that gender differences in preferences over wage-linked firm attributes may be one force contributing to diverging worker-firm allocations and earnings profiles over the life cycle. The next section quantifies the broader preference channel through counterfactual exercises.

Table 4: Preferences over firm characteristics, net of pay

Worker type:	Low-wage		Mid-wage		High-wage	
Gender:	F	M	F	M	F	M
Working hours						
Age 25-30	0.55	0.33	0.47	0.49	-0.29	0.47
Age 31-35	0.46	0.20	-0.03	0.09	-0.12	0.38
Age 36-40	0.56	0.25	-0.38	0.23	0.50	0.45
Age 41-45	0.45	0.28	0.38	0.45	0.62	0.59
Age 45+	0.18	0.25	0.41	0.46	0.60	0.69
Working beyond scheduled hours						
Age 25-30	0.45	0.33	0.43	0.44	-0.06	0.44
Age 31-35	0.50	0.19	0.14	0.16	0.08	0.41
Age 36-40	0.45	0.29	-0.11	0.25	0.53	0.33
Age 41-45	0.52	0.28	0.41	0.36	0.37	0.41
Age 45+	0.31	0.23	0.44	0.40	0.21	0.49

Notes: The table shows correlations between worker preferences and firm-class non-wage attributes, using linked data from DADS and the 2016/2019 CT-RPS surveys. CT-RPS responses are linked to firms in the estimation sample and then aggregated to firm classes k. *Working hours* measures average weekly working hours at the class level. *Working beyond scheduled hours* is the class-level share of linked firms in which at least one respondent reports working beyond scheduled hours either "sometimes", "often", or "every day". Reported entries are, for each gender g, worker type l, and age group, the correlation across firm classes between residualized preferences and residualized class attributes, where both preference terms γ_{glxk} and attributes are residualized on μ_{glxk} within (l, age) cells. The CT-RPS linkage covers about 2% of firms per class.

Gender differences in wages and worker-firm allocations. Figure B2 compares estimated wages and stationary worker-firm allocations between women and men. Panel (a) shows that estimated wages are strongly positively correlated across genders within each (l, k, x) cell, indicating that higher-paying firm classes tend to pay more both women and men. At the same time, the distance from the 45-degree line widens in the upper part of the distribution, showing that gender wage differences are larger for higher-wage worker types in higher-paying firm classes. Panel (b) reveals clear gender differences in worker-firm allocations: many class-specific probabilities lie away from the 45-degree line, and men are almost systematically more likely than women to be allocated to the top-paying firm classes.

6 Worker Sorting and the Gender Wage Gap

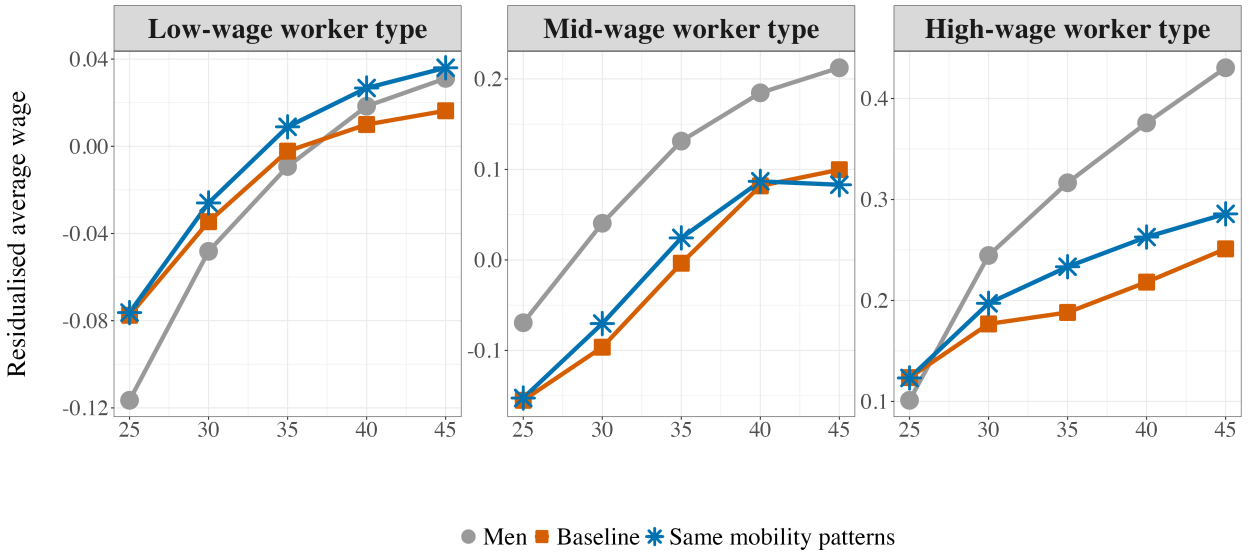
Two patterns documented so far are well known: gender wage profiles diverge over the life cycle, and women are less likely than men to work in the highest-paying firms. The previous section adds a third, novel fact: workers do not value firms similarly, and these differences in valuations arise both within gender and across genders. These three facts thus suggest that lifecycle wage divergence may reflect not only differences in pay conditional on a worker-firm match, but also differences in how women and men move across firms over time. This section aims to quantify the contribution of mobility-driven sorting to lifecycle wage inequality and to assess which mobility channels matter most. The key counterfactual is straightforward: what would women's wage profiles look like if they faced the same mobility patterns as men?

Counterfactual exercise. I simulate cohorts of workers observed at age 25 and follow them for twenty years under the estimated model. The simulation starts from the estimated distribution of initial worker-firm matches and then traces workers forward quarter by quarter using the estimated transition probabilities across firm classes and non-employment. Throughout, the wage schedule μ_{lkx} is held fixed, where l denotes worker type, k firm class, and x the age-gender-tenure cell. Age and tenure evolve mechanically over the simulated career, so differences in wage paths arise from differences in employment states and firm allocations over the life cycle. Wages are model-implied wages residualised on 3-digit occupation and year effects, where the residualisation uses coefficients estimated on the female sample and applied to both men and women. Workers earn no wage while non-employed.

I begin by comparing baseline wage profiles with a counterfactual in which women face the full set of male mobility parameters. Women retain their own estimated wage schedule in each worker-firm match, so any change in wages reflects reallocation rather than changes in pay conditional on a given state. I then study a sequence of cumulative counterfactuals that progressively equalise offer rates while in employment, firm valuations, the incumbent wedge, exit rates, and re-entry transition frequencies. Because the counterfactuals are cumulative, their quantitative contributions depend on the order in which the margins are equalised. I order them to mirror the structure of the estimated mobility process, so the sequence remains

informative about which channels account for the largest changes in simulated wage profiles.[16] I finally consider an additional counterfactual that equalises wage growth along the age-tenure dimension, asking how mobility-driven sorting interacts with gender differences in residual wage progression over the life cycle.

Figure 3: Women's wage profiles under male mobility patterns



Notes: The figure compares simulated wage profiles for men and women by worker type. The grey line reports men's baseline profile, the orange line women's baseline profile, and the blue line the counterfactual profile women would experience if they faced the same mobility parameters as men. Each panel corresponds to a latent worker type. The plotted points are average model-implied wages at ages 25, 30, 35, 40, and 45 from cohorts simulated quarter by quarter from age 25 under the estimated model, starting from the estimated initial distribution of worker-firm matches. Wages are residualised on 3-digit occupation and year effects using coefficients estimated on the female sample and applied to both genders. Reported series are averages across 100 simulation repetitions. The corresponding confidence intervals are reported in Appendix Table X.

Equalizing mobility patterns. Figure 3 compares the baseline wage profiles of women and men to the counterfactual profile women would experience if they faced the same mobility patterns as men. Baseline lifecycle wage gaps differ sharply across worker types. Among low-wage workers, the gap is small throughout, moving from −3.9 log points at age 25 to 1.5 log points at age 45. Among mid-wage workers, it is already substantial at labor-market

16. Appendix Figure X reports the corresponding counterfactuals for each mobility channel separately.

entry, rising from 8.6 log points at age 25 to 13.7 log points at age 30 and remaining around 10 to 11 log points thereafter. Among high-wage workers, the gap widens much more steeply over the life cycle, from −2.2 log points at age 25 to 6.8 log points at age 30, 12.8 log points at age 35, and 17.9 log points at age 45.

Equalizing mobility raises women's wage profiles in all worker types, but the magnitude differs sharply across them. The effects are modest at the bottom of the wage distribution and much larger in the mid- and high-wage types. The contrast is especially sharp at the top. Among high-wage workers, the baseline female profile exhibits a marked mid-career flattening around age 35. When women instead move across firms and employment states as men do, this flattening is substantially attenuated: wages continue to rise through the mid-thirties, and the lifecycle gap opens much more gradually. This pattern is consistent with one mechanism through which the so-called child penalty may emerge, namely that gender differences in mobility and firm allocation become especially consequential precisely at mid-career stages.

I summarize the lifecycle effect by the share of the integrated gender wage gap eliminated by the counterfactual, $1 - \frac{\int \widehat{Gap}^{cf}(a)\,da}{\int \widehat{Gap}^{base}(a)\,da}$, where a indexes age over the simulated life cycle. Equalizing mobility patterns reduces the lifecycle gap by 11% among mid-wage workers and by 29% among high-wage workers. This summary statistic is less informative for low-wage workers, because the lifecycle gap is small and partly offsets itself over age. Quantitatively, then, mobility and reallocation across firms and employment states matter little at the bottom, modestly in the middle, and substantially at the top of the wage distribution.

Contribution of mobility channels. Because the lifecycle gap is both steepest and most responsive to the mobility counterfactual among high-wage workers, I focus the remainder of the analysis on this group.[17] Figure 4 traces the high-wage counterfactual as women's mobility parameters are progressively set equal to the male profile, one channel at a time.

First, equalising offer arrival rates while in employment has virtually no effect and, if anything, slightly widens the lifecycle gap. Panel (a) shows that the counterfactual profile lies close to the baseline one and increases the integrated high-wage gender wage gap by about 10%

17. Appendix Figures X and Y report the results for low-wage and mid-wage types, respectively.

relative to baseline. Although unexpected, this result resonates with a recent correspondence study conducted by the French *Institut des Politiques Publiques*, which found significantly higher callback rates for women in executive roles.[18] This comparison is particularly relevant here because high-wage workers are disproportionately represented in managerial occupations. To the extent that callback rates reflect actual job offers, the results presented in Panel (a) align with the findings of the *Institut des Politiques Publiques*. In addition, the sample period spanning from 2015 to 2019 coincides with increased awareness and action towards gender balance, particularly within the boards of high-paying companies. This counterfactual analysis could capture some of these contemporary trends.

Panels (b) and (c) show that the main gap reduction begins once women face male firm valuations, holding offer arrival rates fixed at the male profile. In panel (b), adding male firm valuations reduces the integrated high-wage lifecycle gap by about 16% relative to baseline. The largest pointwise change occurs at age 35, when the gender gap falls by about 2.8 log points, from 12.8 to 10.0 log points. Adding the incumbent wedge in panel (c) contributes little further, lowering the integrated gap by only about 1 additional percentage point. These results suggest that the job-to-job differences that matter for the high-wage gender gap are driven mainly by differences in how women and men value firms, rather than by differences in employed offer arrival rates or attachment to the incumbent firm.

Panels (d) and (e) shift attention from job-to-job mobility to moves into and out of employment. Equalising exit rates in panel (d) contributes little beyond the previous counterfactual: the integrated high-wage lifecycle gap falls by only 1.5 additional percentage points relative to panel (c). The largest pointwise reduction again occurs at age 35, when the gap narrows by about 3.3 log points relative to baseline. By contrast, equalising re-entry rates in panel (e) produces a much larger additional compression. The integrated gap falls by 29% relative to baseline, about 13 percentage points more than in panel (d), with the largest incremental effect appearing at age 45. These results suggest that, among high-wage workers, mobility into non-employment is of limited importance for the lifecycle wage gap, whereas re-entry patterns after non-employment are quantitatively central. Because re-entry rates

18. (Note IPP n°67) Discrimination à l'embauche selon le sexe: les enseignements d'un testing de grande ampleur.

are left unrestricted in the model, this channel cannot be interpreted as a pure opportunity measure and may also reflect gender differences in preferences over destination firms.

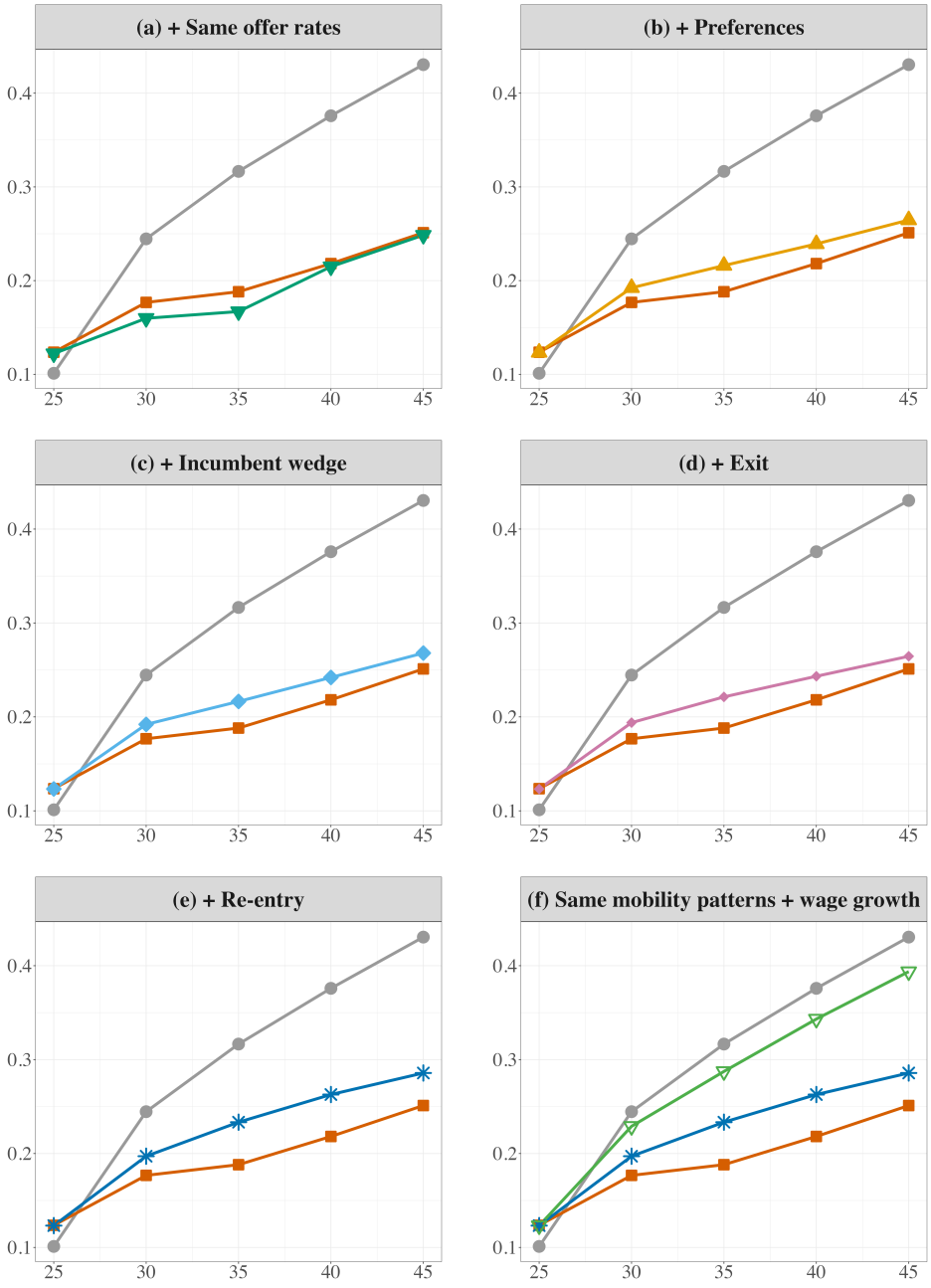
To sum up, the counterfactuals indicate that the sorting component of the high-wage lifecycle gap is shaped mainly by how women value outside firms and by their re-entry patterns after non-employment, rather than by employed offer arrival rates or separations. The fact that the compression is strongest around the mid-thirties and less pronounced by age 45 also suggests that firm reallocation becomes less wage-relevant later in the career, pointing to a role for differential wage growth conditional on worker type, firm class, and age-tenure stage, even after netting out 3-digit occupation and calendar-year effects.

Mobility and residual wage growth. Mobility accounts for 29% of the high-wage lifecycle gap. Panel (f) of Figure 4 adds a final counterfactual that equalises women's residual wage growth to the male profile after equalising mobility. Formally, for each firm class k, worker type l, tenure group tn, and age group e, $\mu^{cf,F}_{lkx(tn,e)} = \mu^{F}_{lkx(tn,1)} + \left(\mu^{M}_{lkx(tn,e)} - \mu^{M}_{lkx(tn,1)}\right)$, so women retain their own wage level at labor-market entry within each tenure group, but inherit the male residual wage-growth increment over the age-tenure dimension.

Because mobility channels are already fixed at the male profile, the additional compression in panel (f) reflects residual wage progression rather than further reallocation across firms or employment states. Quantitatively, the effect is large. The high-wage lifecycle gap falls from 17.9 log points at baseline to 14.5 log points under equal mobility, and to 3.7 log points once residual wage growth is also equalised. In cumulative terms, the lifecycle gap is reduced by about 80% relative to baseline. Since wages are residualised on common 3-digit occupation and year effects, this counterfactual does not capture broad occupation-year composition. Instead, it points to slower residual pay progression for high-wage women even after conditioning on worker type, firm class, age-tenure stage, and common 3-digit occupation and calendar-year effects, consistent with differences in promotions, ranks, responsibilities, or other forms of career advancement not captured by observed occupation codes.

The remaining one-fifth of the integrated lifecycle gap reflects persistent level differences in residual pay even after conditioning on worker type, firm class, age-tenure stage, and common 3-digit occupation and calendar-year effects.

Figure 4: High-type women's wage profiles under male mobility patterns



Notes: The figure reports cumulative counterfactual wage profiles for the high-wage worker type. The grey line reports men's baseline profile (circles) and the orange line women's baseline profile (squares). Across panels, one mobility margin is added at a time: employed offer rates in panel (a), firm valuations in panel (b), the incumbent wedge in panel (c), exit rates in panel (d), and re-entry transition frequencies in panel (e). Panel (f) starts from the full mobility counterfactual and additionally equalises women's residual wage growth over the age-tenure dimension to the male profile while preserving women's wage level at entry. The plotted points are average model-implied wages at ages 25, 30, 35, 40, and 45 from cohorts simulated from age 25 under the estimated model. Wages are residualised on 3-digit occupation and year effects using coefficients estimated on the female sample and applied to both genders. Reported series are averages across 100 simulation repetitions. The corresponding confidence intervals are reported in Appendix Table X.

7 Discussion

This paper combines large-scale matched employer-employee data with a flexible model of wages and mobility to study the sorting component of the gender wage gap first documented by Card, Cardoso, and Kline (2016). Beyond showing that firm allocation matters, the analysis identifies which mobility margins drive the sorting component of lifecycle wage inequality and how their importance varies across the wage distribution. By allowing for rich heterogeneity on both sides of the market, including heterogeneity within and across gender, the framework distinguishes the roles of offer arrival rates, revealed firm valuations, transitions into non-employment, re-entry into employment, and residual wage progression in shaping gender differences in lifecycle wage profiles.

The empirical framework builds on the finite-mixture approach of Lentz, Piyapromdee, and Robin (2023) to model mobility across firms, into non-employment, and back into employment. It is more structured than a reduced-form decomposition, but lighter than a full equilibrium search model. Job-to-job transitions are represented as the outcome of destination-specific offer arrival rates and acceptance decisions based on revealed firm valuations, while an incumbent wedge captures systematic reluctance to move and permits clean gender comparisons on the acceptance margin. This structure allows distinguishing between differences in employment opportunities and firm preferences, while remaining flexible enough to capture worker-firm wage complementarities and gender-specific wage and mobility profiles over the life cycle within a latent worker type.

Using French administrative data for the Paris region over 2015–2019, complemented with firm-level information from the *Conditions de Travail* surveys linked to the administrative records, the analysis shows that sorting contributes substantially to lifecycle gender inequality, especially at the top of the wage distribution. Among high-wage workers, reallocating women across firms and employment states as men are would reduce the integrated lifecycle gap by 29%. Within that sorting component, the quantitatively important margin is the one that reflects how women value firms. Equalising firm valuations reduces the integrated gap by about 16%, or 55% of the total sorting effect. Suggestive evidence from the linked CT-RPS data shows that, among high-wage workers, women's estimated preferences are less positively

associated with firm classes characterized by longer working hours and more frequent overtime, even conditional on pay. Re-entry patterns after non-employment account for most of the remaining sorting component. Because re-entry is left unrestricted in the model, however, the preference share is likely a lower bound, since part of what is attributed to re-entry may also reflect gender differences in valuations of destination firms. Beyond sorting, an additional counterfactual shows that differential residual wage growth remains highly consequential at the top. Once both mobility and residual wage growth are equalised, about 80% of the integrated high-wage lifecycle gap disappears. The remaining 20% reflects persistent level differences in residual pay even after conditioning on worker type, firm class, age-tenure stage, and common 3-digit occupation and calendar-year effects.

Some caveats are worth discussing. First, the framework does not consider idiosyncratic shocks to wages and layoff notifications as explanations of observed moves, potentially confounding voluntary and involuntary transitions in the revealed preference argument. The direction of the resulting bias is not obvious, however, because the analysis leverages flows of groups of workers across groups of firms to average out idiosyncrasies and recover a systematic pattern of preferences.

Second, the analysis is not based on a long panel that would allow one to model age dynamics as done, for example, in Barth, Kerr, and Olivetti (2021) or Amano-Patiño, Baron, and Xiao (2021). The shorter horizon limits the extent to which worker heterogeneity can be identified from within-worker wage growth alone. The finite-mixture approach partly addresses this constraint by allowing for latent worker heterogeneity both within and across gender, while interactions between gender, age, and tenure categories provide a tractable way to capture differences in wage profiles and mobility patterns across career stages.

Finally, while sorting based on preferences explains a sizeable share of the residualized gender wage gap, these findings should not be read as dismissing the need for corrective action. Gendered norms often impose gendered roles in caregiving responsibilities. Although we may claim that they reflect preferences, we also know that preferences may internalize prescriptive norms when group identity encourages specific choices (Akerlof and Kranton, 2000), complicating the distinction between preferences, norms, and constraints. If women's choices reflect gender stereotypes, they also blur the meaning of preferences (Bertrand,

2020). The findings of this paper recall proposals to restructure jobs so that a broader range of possibly constrained workers can reach them (Goldin, 2014; Goldin and Katz, 2016; Wasserman, 2022).

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333, http://www.jstor.org/stable/2999586.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* 115 (3): 715–753, http://www.jstor.org/stable/2586894.

Amano-Patiño, Noriko, Tatiana Baron, and Pengpeng Xiao. 2021. "Human Capital Accumulation, Equilibrium Wage-Setting and the Life-Cycle Gender Pay Gap." Working Paper 2010, Cambridge Working Papers in Economics.

Angelici, Marta, and Paola Profeta. 2020. "Smart-Working: Work Flexibility Without Constraints." CESifo Working Paper 8165, CESifo, https://ssrn.com/abstract=3556304orhttp://dx.doi.org/10.2139/ssrn.3556304.

Arcidiacono, Peter, Attila Gyetvai, Arnaud Maurel, and Ekaterina Jardim. 2025. "Identification and Estimation of Continuous-Time Job Search Models with Preference Shocks." IZA Discussion Paper 18309, Institute of Labor Economics (IZA), Bonn, https://www.iza.org/publications/dp18309.

Bagger, Jesper, and Rasmus Lentz. 2019. "An Empirical Model of Wage Dispersion with Sorting." *The Review of Economic Studies* 86 (1): 153–190. 10.1093/restud/rdy022.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti. 2021. "The dynamics of gender earnings differentials: Evidence from establishment data." *European Economic Review* 134 103713. https://doi.org/10.1016/j.euroecorev.2021.103713.

Bertrand, Marianne. 2020. "Gender in the Twenty-First Century." *AEA Papers and Proceedings* 110 1–24. 10.1257/pandp.20201126.

Blau, Francine D. 1977. *Equal Pay in the Office*. Washington, DC: Lexington Books.

Blau, Francine D., Peter Brummund, and Albert Yung-Hsu Liu. 2012. "Trends in Occupational Segregation by Gender 1970–2009: Adjusting for the Impact of Changes in the Occupational Coding System." *Demography* 50 (2): 471–492. 10.1007/s13524-012-0151-7.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789–865. 10.1257/jel.20160995.

Bonhomme, Stéphane, and Angela Denis. 2025. "Fixed Effects and Beyond. Bias Reduction, Groups, Shrinkage and Factors in Panel Data." Working Papers 2526, Banco de Espana.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2023. "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics* 41 (2): 291–322. 10.1086/720009.

Bonhomme, Stéphane, and Grégory Jolivet. 2009. "The Pervasive Absence of Compensating Differentials." *Journal of Applied Econometrics* 24 (5): 763–795, http://www.jstor.org/stable/25608759.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica* 87 (3): 699–739. https://doi.org/10.3982/ECTA15722.

Borovickova, Katarina, and Robert Shimer. 2025. "Assortative Matching and Wages: The Role of Selection." Working Paper.

Bowlus, Audra J. 1997. "A Search Interpretation of Male-Female Wage Differentials." *Journal of Labor Economics* 15 (4): 625–657. 10.1086/209840.

Bradley, Ralph Allan, and Milton E. Terry. 1952. "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons." *Biometrika* 39 (3/4): 324–345, http://www.jstor.org/stable/2334029.

Braun, Christine, and Ana Figueiredo. 2025. "Labor Market Beliefs and the Gender Wage Gap." Working Paper.

Bruns, Benjamin. 2019. "Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap." *American Economic Journal: Applied Economics* 11 (2): 74–113. 10.1257/app.20160664.

Caldwell, Sydnee, and Oren Danieli. 2024. "Outside Options in the Labour Market." *The Review of Economic Studies* rdae006. 10.1093/restud/rdae006.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. " Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women *." *The Quarterly Journal of Economics* 131 (2): 633–686. 10.1093/qje/qjv038.

Casarico, Alessandra, and Salvatore Lattanzio. 2024. "What Firms Do: Gender Inequality in Linked Employer-Employee Data." *Journal of Labor Economics* 42 (2): 325–355. 10.1086/723177.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar. 2023. "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab*." *The Quarterly Journal of Economics* 138 (4): 2069–2126. 10.1093/qje/qjad017.

Coudin, Elise, Sophie Maillard, and Maxime To. 2018. "Family, Firms and the Gender Wage Gap in France." IFS Working Papers W18/01, Institute for Fiscal Studies, https://ideas.repec.org/p/ifs/ifsewp/18-01.html.

Cruz, Gabriel, and Tomás Rau. 2022. "The effects of equal pay laws on firm pay premiums: Evidence from Chile." *Labour Economics* 75 102135. https://doi.org/10.1016/j.labeco.2022.102135.

Davidson, Roger R., and Robert J. Beaver. 1977. "On Extending the Bradley-Terry Model to Incorporate Within-Pair Order Effects." *Biometrics* 33 (4): 693–702, http://www.jstor.org/stable/2529467.

Del Bono, Emilia, and Daniela Vuri. 2011. "Job mobility and the gender wage gap in Italy." *Labour Economics* 18 (1): 130–142. https://doi.org/10.1016/j.labeco.2010.06.002.

Flinn, C., and J. Heckman. 1982. "New methods for analyzing structural models of labor force dynamics." *Journal of Econometrics* 18 (1): 115–168. https://doi.org/10.1016/0304-4076(82)90097-5.

Fluchtmann, Jonas, Anita M. Glenny, Nikolaj A. Harmon, and Jonas Maibom. 2024. "The Gender Application Gap: Do Men and Women Apply for the Same Jobs?" *American Economic Journal: Economic Policy* 16 (2): 182–219. 10.1257/pol.20210607.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119. 10.1257/aer.104.4.1091.

Goldin, Claudia, and Lawrence F. Katz. 2016. "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation." *Journal of Labor Economics* 34 (3): 705–746. 10.1086/685505.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth. 2017. "The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census." *American Economic Review* 107 (5): 110–14. 10.1257/aer.p20171065.

Heckman, J., and B. Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52 (2): 271–320, http://www.jstor.org/stable/1911491.

Hirschman, Daniel. 2022. "Controlling for What? Movements, Measures, and Meanings in the US Gender Wage Gap Debate." *History of Political Economy* 54 (S1): 221–257. 10.1215/00182702-10085710.

Hunter, David R. 2004. "MM algorithms for generalized Bradley-Terry models." *The Annals of Statistics* 32 (1): 384 – 406. 10.1214/aos/1079120141.

Hunter, David R, and Kenneth Lange. 2004. "A Tutorial on MM Algorithms." *The American Statistician* 58 (1): 30–37. 10.1198/0003130042836.

Jain, A.K., and J.V. Moreau. 1987. "Bootstrap technique in cluster analysis." *Pattern Recognition* 20 (5): 547–568. https://doi.org/10.1016/0031-3203(87)90081-1.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. "Systemic Discrimination Among Large U.S. Employers*." *The Quarterly Journal of Economics* 137 (4): 1963–2036. 10.1093/qje/qjac024.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading off Commute against Wage*." *The Quarterly Journal of Economics* 136 (1): 381–426. 10.1093/qje/qjaa033.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin. 2023. "The Anatomy of Sorting—Evidence From Danish Data." *Econometrica* 91 (6): 2409–2455. https://doi.org/10.3982/ECTA16425.

Li, Jiang, Benoit Dostie, and Gaëlle Simard-Duplain. 2020. "What Is the Role of Firm-Specific Pay Policies on the Gender Earnings Gap in Canada?." IZA Discussion Paper 13907, IZA, https://docs.iza.org/dp13907.pdf.

Loprest, Pamela J. 1992. "Gender Differences in Wage Growth and Job Mobility." *The American Economic Review* 82 (2): 526–532, http://www.jstor.org/stable/2117456.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12): 3722–59. 10.1257/aer.20161500.

Morchio, Iacopo, and Christian Moser. 2023. "The Gender Pay Gap: Micro Sources and Macro Consequences." Institute working paper 78, Federal Reserve Bank of Minneapolis. Opportunity and Inclusive Growth Institute. https://doi.org/10.21034/iwp.78.

Neumark, David, Roy J. Bank, and Kyle D. Van Nort. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *The Quarterly Journal of Economics* 111 (3): 915–941, http://www.jstor.org/stable/2946676.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics* 8 (1): 405–434. 10.1146/annurev-economics-080614-115329.

Palladino, Marco G., Antoine Berthou, Alexander Hijzen et al. 2025b. "Firms and the Gender Wage Gap: A Comparison of Eleven Countries." Working Paper 2025-24, Federal Reserve Bank of Chicago. https://doi.org/10.2139/ssrn.5943934.

Palladino, Marco G., Alexandra Roulet, and Mark Stabile. 2025c. "Narrowing industry wage premiums and the decline in the gender wage gap." *Labour Economics* 94 102693. https://doi.org/10.1016/j.labeco.2025.102693.

Palladino, Marco Guido, Damien Babet, and Olivier Godechot. 2025a. "In the Land of AKM: Explaining the Dynamics of Wage Inequality in France." *Journal of Labor Economics* 0 (ja): null. 10.1086/740187.

Petrongolo, Barbara, and Maddalena Ronchi. 2020. "Gender gaps and the structure of local labor markets." *Labour Economics* 64 101819. https://doi.org/10.1016/j.labeco.2020.101819, European Association of Labour Economists, 31st annual conference, Uppsala Sweden, 19-21 September 2019.

Samuelson, Larry, and Jakub Steiner. 2025. "Robust Latent Data Representations." Working Paper.

Sorkin, Isaac. 2017. "The Role of Firms in Gender Earnings Inequality: Evidence from the United States." *American Economic Review: Papers and Proceedings* 107 (5): 384–87. 10.1257/aer.p20171015.

Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference*." *The Quarterly Journal of Economics* 133 (3): 1331–1393. 10.1093/qje/qjy001.

Taber, Christopher, and Rune Vejlin. 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica* 88 (3): 1031–1069. https://doi.org/10.3982/ECTA14441.

Wasserman, Melanie. 2022. "Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents." *The Review of Economic Studies* 90 (3): 1535–1568. 10.1093/restud/rdac042.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender*." *The Quarterly Journal of Economics* 133 (1): 457–507. 10.1093/qje/qjx035.

Xiao, Pengpeng. 2023. "Equilibrium Sorting and the Gender Wage Gap." Working Paper 144, VATT Institute for Economic Research. https://urn.fi/URN:ISBN:978-952-274-278-0.

Online Appendix

A Appendix Tables

Table A1: Firms' characteristics across selection samples

	Sample 1	Sample 2	Sample 3 (analysis)
N firms	589,019	148,967	126,587
N active years	3	5	5
N employees	21	72	83
N employees in Ile de France	12	41	48
Share of women	70%	45%	44%
Share of women that are managers	6%	16%	17%
Share of women in the board	41%	42%	42%
Share of men	30%	55%	56%
Share of men that are managers	16%	20%	21%
Share of men in the board	59%	58%	58%
Average earnings (EUR)	14,287	27,033	27,799
Median earnings (EUR)	13,559	24,752	25,350
Average hours	793	1,312	1,329
Median hours	817	1,394	1,423
Average hourly wages (EUR)	17	20	20
Median hourly wages (EUR)	16	18	18
Share with part-time contracts	7%	17%	16%
Share part-time and females	4%	10%	10%

Notes: The table compares firms' characteristics across three different selection samples. Sample 1 includes firms with at least one full-time job in Ile de France in 2015. Sample 2 includes those hiring both genders and being active for five consecutive years. Sample 3, used in the analysis, includes those firms where I observe workers selected for the analysis.

Table A2: Workers' characteristics across selection steps

	Earnings	Hours	Age	Tenure	% Part-Time	% Managers	N
	Women						
Main job	31,213	1,417	40	6	18%	29%	1,555,053
Ile de France only	32,121	1,449	40	6	18%	29%	1,426,866
30+ days contract	32,196	1,454	40	6	18%	29%	1,415,043
+wages, +hours	32,220	1,456	40	6	18%	29%	1,411,587
Never in Agriculture	32,221	1,456	40	6	18%	29%	1,411,026
Part-time & Full-time only	30,816	1,491	40	6	19%	24%	1,214,070
Aged 25-55	30,816	1,491	40	6	19%	24%	1,214,070
Never in seasonal/internship/domicile	31,759	1,520	40	7	18%	24%	1,130,653
Only in selected firms	32,565	1,552	40	7	18%	24%	1,066,931
	Men						
	Earnings	Hours	Age	Tenure	% Part-Time	% Managers	N
Main job	38,676	1,483	40	6	8%	35%	1,702,116
Ile de France only	40,118	1,526	40	6	8%	34%	1,528,015
30+ days contract	40,256	1,533	40	6	8%	34%	1,512,966
+wages, +hours	40,275	1,535	40	6	8%	34%	1,509,837
Never in Agriculture	40,281	1,535	40	6	8%	34%	1,508,658
Part-time & Full-time only	37,520	1,578	40	6	8%	28%	1,242,338
Aged 25-55	37,520	1,578	40	6	8%	28%	1,242,338
Never in seasonal/internship/domicile	38,712	1,607	40	7	7%	28%	1,162,855
Only in selected firms	39,940	1,644	41	7	6%	29%	1,085,686

Notes: This table reports workers' characteristics across sequential sample-selection steps using DADS-Postes data. The top panel reports women; the bottom panel reports men. *Main job* includes workers employed in 2015Q1 in Ile-de-France with at least one spell classified as a main job. *Ile-de-France only* keeps workers whose observed spells are all in Ile-de-France. *30+ days contract* keeps workers whose observed spells are all in Ile-de-France. *+wages, +hours* keeps spells with strictly positive wages and hours. *Never in agriculture* excludes workers with any spell in the agricultural sector. *Part-time & full-time only* keeps workers holding only PT or FT jobs. *Aged 25–55* applies the age restriction. *Never in seasonal/internship/domicile* excludes workers ever observed in those contract types. *Only in selected firms* keeps workers who, when employed, are observed only in the selected-firm sample (firms with at least one full-time job in Ile-de-France in 2015, employing both genders, and active in every year from 2015 to 2019).

Table A3: Sample description by gender, age, and tenure

	Women									
Tenure:	Short					Long				
Age:	25-30	31-35	36-40	41-45	45+	25-30	31-35	36-40	41-45	45+
Mean Annual Earnings	23,527	25,792	27,563	29,241	29,178	28,432	31,259	34,550	37,161	38,112
Mean Hours	1,416	1,429	1,428	1,449	1,445	1,625	1,635	1,657	1,684	1,699
Share Part-time	11%	14%	18%	20%	20%	10%	14%	18%	19%	17%
Mean Age	28	32	37	42	49	28	32	37	42	49
Share doing JTJ	12%	8%	7%	6%	6%	16%	13%	10%	8%	8%
Share doing E-NE	24%	21%	19%	17%	17%	29%	25%	21%	16%	17%
Share doing NE-E	17%	16%	14%	11%	10%	5%	3%	3%	2%	2%
Mean job spell quarters	3	3	3	3	4	7	8	9	9	13
N workers	117,588	149,331	128,037	107,927	136,647	157,685	333,442	350,914	356,240	453,279
	Men									
Tenure:	Short					Long				
Age:	25-30	31-35	36-40	41-45	45+	25-30	31-35	36-40	41-45	45+
Mean Annual Earnings	26,687	29,972	32,486	35,050	37,986	33,183	37,280	41,276	44,731	48,308
Mean Hours	1,509	1,536	1,550	1,553	1,560	1,716	1,743	1,763	1,775	1,783
Share Part-time	7%	6%	6%	7%	7%	5%	4%	4%	4%	4%
Mean Age	28	32	37	42	49	28	32	37	42	49
Share doing JTJ	9%	7%	6%	6%	5%	14%	12%	10%	7%	7%
Share doing E-NE	25%	22%	21%	20%	19%	26%	22%	19%	17%	20%
Share doing NE-E	12%	9%	8%	8%	8%	3%	2%	2%	1%	1%
Mean job spell quarters	3	3	3	3	4	7	8	9	9	13
N workers	107,112	140,388	124,796	114,994	150,813	149,637	326,980	357,975	374,933	470,665

Notes: This table reports descriptive statistics for the analysis sample, by gender-age-tenure cells. The sample period is 2015Q1–2019Q3. Mobility measures are the share of workers who experience at least one transition of the indicated type during the sample period. *JTJ* denotes job-to-job transitions; *E* and *NE* denote employment and non-employment, respectively. Earnings and hours are reported as annual averages. *N workers* is the number of worker-cell observations over 2015Q1–2019Q3. Because workers can change age and tenure categories over time, the same worker may appear in multiple columns.

Table A4: Wages, hours, and mobility by worker type, age, gender, and tenure

Tenure:	Short						Long					
Worker type:	Low-wage		Mid-wage		High-wage		Low-wage		Mid-wage		High-wage	
Gender:	F	M	F	M	F	M	F	M	F	M	F	M
Hourly wage (EUR)												
Age 25-30	11.29	12.51	11.61	12.02	12.45	13.99	13.24	14.78	14.19	14.48	15.23	18.04
Age 31-35	9.43	10.74	9.83	11.19	11.70	14.19	13.55	15.15	14.79	15.69	16.79	20.46
Age 36-40	9.28	10.23	10.06	11.52	12.37	15.34	14.99	15.95	16.43	18.02	19.70	24.48
Age 41-45	9.74	10.52	11.09	12.06	13.66	17.79	15.91	16.66	18.13	19.75	22.57	28.76
Age 45+	9.01	9.40	9.97	11.03	12.56	16.85	16.11	16.84	18.69	20.83	24.51	32.25
Hours												
Age 25-30	1,106	1,225	1,117	1,100	1,046	1,117	1,381	1,470	1,348	1,384	1,296	1,385
Age 31-35	878	1,046	884	938	837	928	1,297	1,431	1,299	1,351	1,225	1,314
Age 36-40	850	1,006	859	886	797	871	1,359	1,489	1,381	1,426	1,291	1,359
Age 41-45	912	1,013	902	875	814	851	1,436	1,538	1,466	1,478	1,371	1,407
Age 45+	860	902	831	790	725	721	1,488	1,557	1,513	1,517	1,411	1,391
Share doing JTJ												
Age 25-30	0.12	0.09	0.14	0.07	0.10	0.10	0.14	0.15	0.20	0.11	0.15	0.17
Age 31-35	0.07	0.07	0.10	0.06	0.08	0.08	0.11	0.13	0.14	0.10	0.14	0.15
Age 36-40	0.06	0.07	0.07	0.05	0.08	0.07	0.09	0.10	0.09	0.07	0.12	0.13
Age 41-45	0.06	0.06	0.06	0.05	0.08	0.06	0.07	0.08	0.07	0.05	0.11	0.10
Age 45+	0.05	0.06	0.05	0.05	0.08	0.05	0.08	0.08	0.07	0.05	0.11	0.10
Share doing E-NE												
Age 25-30	0.25	0.22	0.21	0.28	0.27	0.24	0.27	0.22	0.30	0.29	0.32	0.30
Age 31-35	0.22	0.21	0.18	0.23	0.22	0.22	0.23	0.18	0.26	0.22	0.28	0.27
Age 36-40	0.20	0.20	0.18	0.22	0.21	0.19	0.18	0.15	0.20	0.19	0.24	0.24
Age 41-45	0.17	0.19	0.18	0.20	0.17	0.20	0.14	0.13	0.15	0.17	0.20	0.23
Age 45+	0.17	0.19	0.16	0.19	0.18	0.20	0.15	0.16	0.14	0.18	0.23	0.28
Share doing NE-E												
Age 25-30	0.14	0.12	0.24	0.09	0.16	0.14	0.04	0.04	0.06	0.03	0.05	0.04
Age 31-35	0.13	0.10	0.20	0.07	0.16	0.11	0.03	0.03	0.05	0.01	0.03	0.03
Age 36-40	0.12	0.10	0.15	0.05	0.15	0.10	0.03	0.02	0.03	0.01	0.03	0.02
Age 41-45	0.10	0.10	0.09	0.05	0.14	0.09	0.02	0.02	0.02	0.01	0.03	0.02
Age 45+	0.09	0.11	0.08	0.05	0.13	0.09	0.03	0.02	0.02	0.01	0.03	0.01
N workers in $t = 1$	145,779	143,453	99,368	108,731	109,281	101,755	275,539	268,045	219,173	259,217	217,660	204,300
N observations	1,015,544	989,972	739,765	807,448	875,614	826,132	7,410,818	7,240,032	5,631,073	6,551,535	5,663,227	5,294,981

Notes: The table presents ex-post tabulations of workers' wages and mobility rates by predicted latent types and observed age, tenure, and gender categories. Workers are classified into types using the EM algorithm described in Section 3.2. Data relative to unconditional wages correspond to the pooled 2015–2019 sample and represent averages weighted by the estimated posterior probability $p_i\Big(l\Big|\theta^{(m)}, x_{i1}, C\Big)$. Data relative to mobility represents the share of workers doing at least one given transition over the sample period, again weighted by the estimated posterior probability. JTJ stands for Job-To-Job. E stands for Employment. NE stands for Non-Employment. *N workers in* $t = 1$ is the sum of the estimated posterior probability conditional on first-period tenure.

N observations is $\sum_{i=1}^{N} p_i\Big(l \;\Big|\; \theta^{(m)}, x_{i1}, C\Big) \sum_t \mathbb{1}\Big\{x_{it} = x\Big\}$.

Table A5: Share of women by worker type, age, and tenure

Tenure:	Short			Long		
Worker type:	Low-wage	Mid-wage	High-wage	Low-wage	Mid-wage	High-wage
Age						
Age 25-30	0.52	0.50	0.54	0.53	0.44	0.52
Age 31-35	0.50	0.46	0.52	0.52	0.44	0.53
Age 36-40	0.50	0.46	0.52	0.50	0.44	0.53
Age 41-45	0.49	0.48	0.49	0.49	0.46	0.51
Age 45+	0.50	0.48	0.49	0.50	0.48	0.50

Notes: The table shows the female shares by worker type, age, and tenure. It is calculated from the sums of the estimated posterior probability conditional on age and tenure categories.

Table A6: Occupational composition by worker type, age, gender, and tenure

Tenure:	Short						Long					
Worker type:	Low-wage		Mid-wage		High-wage		Low-wage		Mid-wage		High-wage	
Gender:	F	M	F	M	F	M	F	M	F	M	F	M
Share managers												
Age 25-30	0.23	0.37	0.40	0.29	0.18	0.29	0.16	0.32	0.35	0.20	0.15	0.24
Age 31-35	0.22	0.30	0.35	0.31	0.28	0.36	0.18	0.30	0.34	0.21	0.23	0.32
Age 36-40	0.20	0.25	0.27	0.32	0.34	0.40	0.21	0.27	0.27	0.22	0.31	0.38
Age 41-45	0.18	0.25	0.25	0.32	0.36	0.45	0.21	0.26	0.22	0.23	0.35	0.41
Age 45+	0.15	0.23	0.22	0.30	0.34	0.46	0.17	0.23	0.19	0.22	0.34	0.42
Share intermediate managers												
Age 25-30	0.32	0.18	0.25	0.17	0.21	0.20	0.37	0.22	0.29	0.20	0.24	0.22
Age 31-35	0.28	0.18	0.25	0.18	0.23	0.20	0.37	0.22	0.30	0.22	0.27	0.23
Age 36-40	0.25	0.17	0.27	0.19	0.24	0.20	0.34	0.22	0.33	0.23	0.28	0.22
Age 41-45	0.22	0.16	0.28	0.20	0.24	0.18	0.29	0.22	0.37	0.23	0.27	0.21
Age 45+	0.18	0.15	0.28	0.21	0.24	0.18	0.24	0.21	0.35	0.24	0.27	0.21
Share non-manual employees												
Age 25-30	0.40	0.27	0.33	0.30	0.56	0.31	0.41	0.25	0.33	0.25	0.57	0.29
Age 31-35	0.42	0.28	0.36	0.24	0.44	0.23	0.38	0.25	0.32	0.23	0.45	0.22
Age 36-40	0.42	0.30	0.38	0.21	0.36	0.18	0.37	0.24	0.34	0.23	0.37	0.18
Age 41-45	0.43	0.28	0.38	0.21	0.33	0.15	0.39	0.24	0.35	0.23	0.32	0.15
Age 45+	0.45	0.26	0.40	0.18	0.33	0.13	0.42	0.22	0.38	0.19	0.32	0.13
Share manual employees												
Age 25-30	0.05	0.18	0.03	0.24	0.05	0.19	0.05	0.22	0.03	0.35	0.05	0.24
Age 31-35	0.08	0.23	0.05	0.26	0.05	0.20	0.06	0.23	0.04	0.33	0.04	0.22
Age 36-40	0.12	0.28	0.08	0.27	0.05	0.20	0.07	0.26	0.05	0.31	0.04	0.20
Age 41-45	0.17	0.30	0.09	0.26	0.06	0.20	0.11	0.28	0.06	0.31	0.04	0.20
Age 45+	0.22	0.35	0.10	0.29	0.08	0.20	0.16	0.34	0.07	0.35	0.05	0.22

Notes: The table reports ex-post occupational composition by predicted worker type, gender, age group, and tenure status. Workers are assigned to latent types using the EM algorithm described in Section 3.2. For each latent worker type l and observed cell x, each entry reports the posterior-probability-weighted share of worker-year observations in the indicated occupation category over the pooled 2015–2019 sample. The weights are $p_i\left(l \mid \theta^{(m)}, x_{i1}, C\right)$. Occupation groups correspond to the first digit of the PCS classification: managers, intermediate managers, non-manual employees, and manual employees.

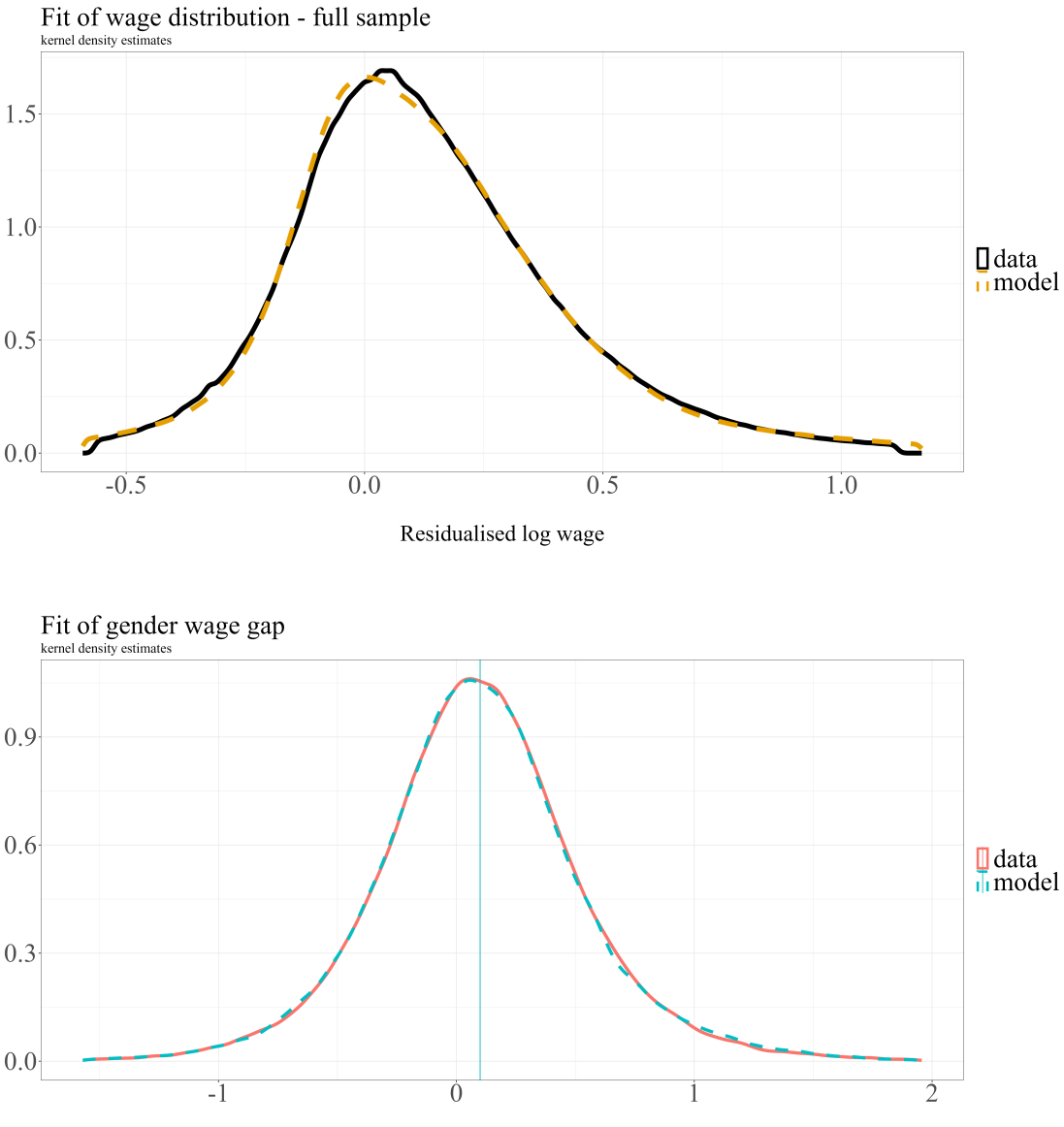
Table A7: Wages and firm characteristics across firm classes

Worker type:	Low-wage		Mid-wage		High-wage	
Gender:	F	M	F	M	F	M
Corr(working hours, wages)						
Age 25-30	0.47	0.37	0.44	0.60	0.56	0.73
Age 31-35	0.52	0.50	0.42	0.70	0.64	0.71
Age 36-40	0.50	0.50	0.41	0.73	0.70	0.64
Age 41-45	0.44	0.48	0.46	0.71	0.65	0.59
Age 45+	0.44	0.49	0.40	0.70	0.61	0.57
Corr(overtime incidence, wages)						
Age 25-30	0.40	0.41	0.27	0.51	0.51	0.60
Age 31-35	0.43	0.52	0.36	0.62	0.59	0.63
Age 36-40	0.42	0.50	0.39	0.69	0.63	0.58
Age 41-45	0.34	0.49	0.54	0.68	0.53	0.57
Age 45+	0.33	0.50	0.53	0.68	0.51	0.56

Notes: The table reports correlations between firm-class wages and firm-class non-wage attributes, using linked data from DADS and the 2016/2019 CT-RPS surveys. CT-RPS responses are linked to firms in the estimation sample and then aggregated to firm classes k. *Working hours* measures average weekly working hours at the class level. *Working beyond scheduled hours* is the class-level share of linked firms in which at least one respondent reports working beyond scheduled hours either "sometimes", "often", or "every day". Reported entries are, for each gender g, worker type l, and age group, the correlation across firm classes between wage components μ_{glxk} and class attributes. The CT-RPS linkage covers about 2% of firms per class.

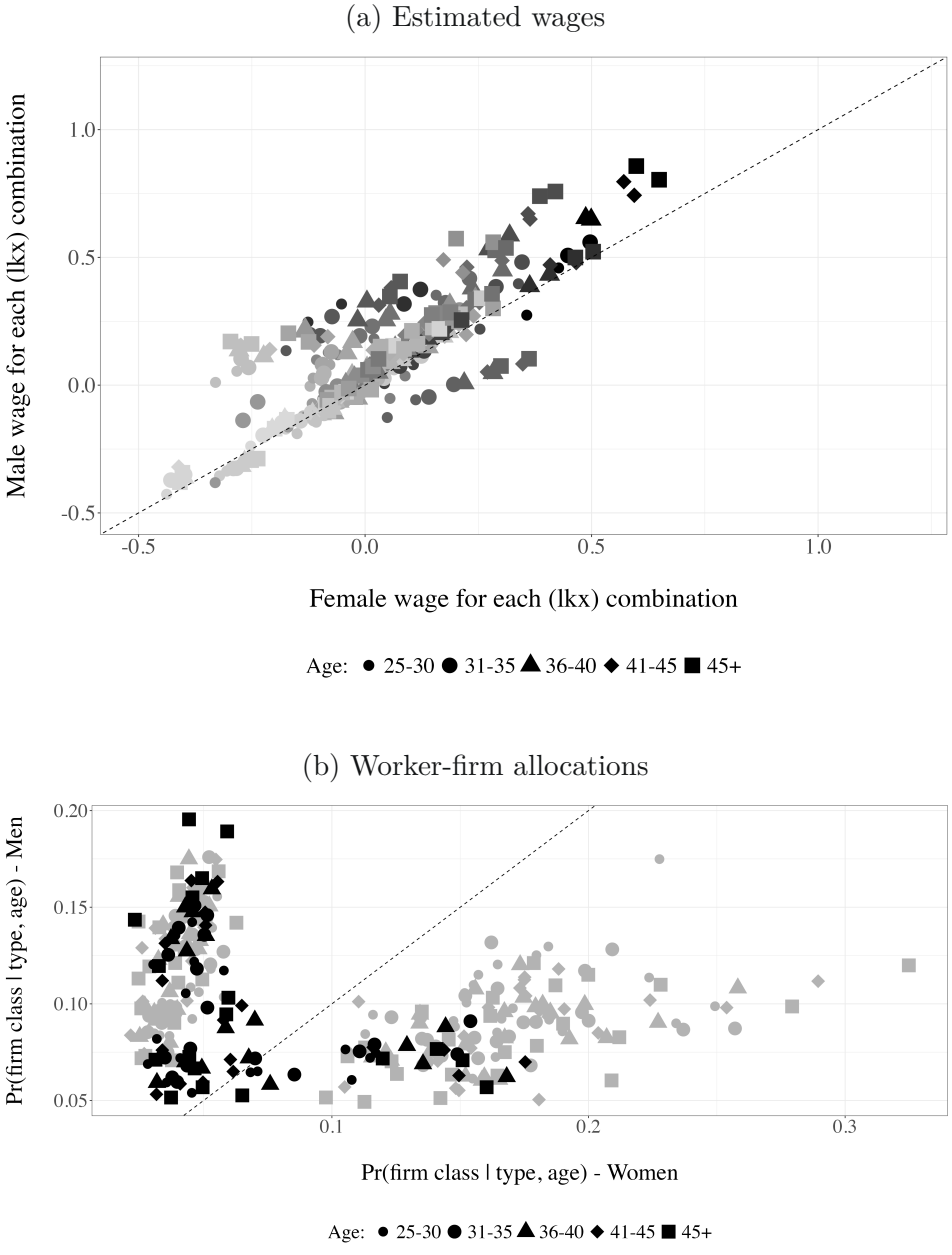
B Appendix Figures

Figure B1: Fit of wage distribution and gender wage gap



Notes: The figure assesses model fit by comparing observed wage distributions with model-implied distributions from data simulated using the estimated parameters. The top panel compares the observed wage distribution (solid line) with the model-implied distribution (dashed line) for the full sample. The bottom panel reports the gender wage gap, computed as the difference between the male and female wage distributions, in the observed data (solid line) and in the model-implied data (dashed line).

Figure B2: Gender differences in estimated wages and worker-firm allocations



Notes: The figure compares estimated wages and stationary worker-firm allocations between women and men. Panel (a) plots the estimated wage component $\hat{\mu}_{lkx}$ for women on the x-axis and for men on the y-axis, for each worker type l, firm class k, and age-tenure cell x. Darker shades correspond to higher-wage worker types in higher-paying firm classes, and marker shapes indicate age groups. Panel (b) plots the stationary allocation probability $\Pr^*(k \mid l, x)$ for women on the x-axis and for men on the y-axis, for the same (l, k, x) combinations. In Panel (b), black markers denote the top three paying firm classes and grey markers denote all other classes. The dashed line in both panels is the 45-degree line.

C K-means algorithm for firm clustering

I partition the J firms into K distinct classes by solving a weighted k-means problem. As input characteristics, I use firms' male and female empirical cumulative distribution functions of log-hourly wages computed over five years and mean female shares, weighting by firms' average sizes. Log wages are residualised on 3-digit occupational and year dummies.

For a given K, I initiate the algorithm with 100 initial guesses and select the classification with the smallest residual sum of squares, i.e. the sum of the squared deviations from each observation and the cluster centroid. Figure C1 shows the curve of the within-class sum of squares for a given number of classes K. I choose $K = 10$, seeking a balance between minimising total intra-class variation and ensuring sufficient observations.

To assess cluster stability, I implement a bootstrap procedure (Jain and Moreau, 1987). For a fixed K, I repeatedly generate firm-level datasets through random sampling with replacement, maintaining the original dataset's size. I then cluster the newly sampled data. Using the Jaccard similarity, I identify the most similar boostrap cluster for each cluster in the original classification, repeating this process 100 times. Table C1 reports, for each firm class, the average Jaccard similarity computed across these repetitions, showing the overall stability of firm classes.

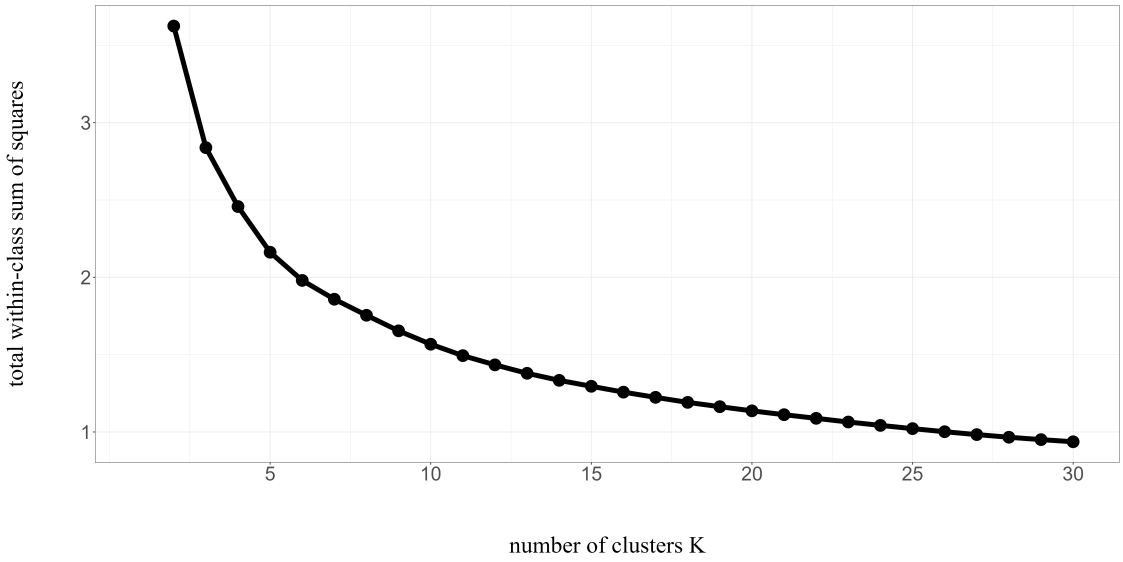
To explore the robustness of the gender wage gap decomposition results, I vary the number of clusters K and re-estimate the full model and its counterfactuals. I report two alternatives, with $K = 5$ and $K = 15$, in Table **??** and Table **??**, respectively. Results are unaffected.

Table C1: Firm clusters stability

Firm class:	1	2	3	4	5	6	7	8	9	10
Avg Jaccard similarity	0.86	0.80	0.68	0.69	0.56	0.68	0.52	0.62	0.56	0.73

Notes: For each firm class $k \in \{1, \ldots, K\}$, the table reports the average Jaccard similarity computed across 100 bootstrapped repetitions. The labels of firm classes are consistent with the ones reported in the main text.

Figure C1: Within firm-class variation by number of classes



Notes: The figure plots the total within class sum of squares (y-axis), rescaled by the total number of firms, against increasing numbers of firm clusters K (x-axis).

D EM and MM algorithm

This section describes in detail the algorithm of Lentz, Piyapromdee, and Robin (2023) to classify workers into a finite number of groups and to estimate the parameters of interest. The expectation-maximization (EM) algorithm is an iterative method used to numerically find local maximum likelihood parameters in statistical models that commonly involve latent variables in addition to unknown parameters and observed data points. It is desirable and practical for maximum likelihood estimation because, at each iteration, it consistently increases the likelihood by maximizing a simple surrogate function for the log-likelihood. The EM algorithm is a particular case of a more general class of optimization algorithms, minorization-maximization (MM) algorithms, that exploit concavity in finding a surrogate function for maximization (Hunter and Lange, 2004).

The basic idea of the MM algorithm is to look for a minorizing function that facilitates the maximization step. Let the real-valued objective function be $f(\theta)$. A real-valued function $g(\theta|\theta^{(s)})$ is said to be minorizing $f(\theta)$ if $g(\theta|\theta^{(s)}) \leq f(\theta)\ \forall\theta$ and $g(\theta^{(s)}|\theta^{(s)}) = f(\theta^{(s)})$. $\theta^{(s)}$ is the parameter vector obtained at the current iteration, and $g(.)$ is the surrogate function being maximized in the M-step of the algorithm. If $\theta^{(s+1)}$ is the local maximizer of $g(\theta|\theta^{s})$, then $f(\theta^{(s+1)}) \geq f(\theta^{s})$.[19]

The log-likelihood relative to the likelihood in equation 1 is not linear in the mobility components of the parameter vector θ. To simplify the maximization step, I build on the surrogate function proposed by Lentz, Piyapromdee, and Robin (2023) and adapt it to accommodate the inertia parameter. In this Appendix, I present the resulting surrogate log-likelihood and derive the first-order conditions for the parameters of interest.

As it is a local estimation, I initiate the full estimation with 20 random guesses for the parameters of interest. For the computation, I parallelize each repetition across different CPUs. I select the repetition with the highest likelihood. The statistical software used is R.

19. Indeed, $g(\theta^{(s+1)}|\theta^{(s)}) \geq g(\theta^{(s)}|\theta^{(s)})$ by definition. Together with the definition of the function $g(\theta|\theta^{(s)})$, that is $g(\theta|\theta^{(s)}) \leq f(\theta)\ \forall\theta$ and $g(\theta^{(s)}|\theta^{(s)}) = f(\theta^{(s)})$, it is straightforward to see that the following inequality is true.

$$f(\theta^{(s+1)}) = g(\theta^{(s+1)}|\,\theta^{(s)}) + f(\theta^{(s+1)}) - g(\theta^{(s+1)}|\,\theta^{(s)}) \geq g(\theta^{(s)}|\,\theta^{(s)}) + f(\theta^{(s)}) - g(\theta^{(s)}|\,\theta^{(s)}) = f(\theta^{(s)})$$

D.1 Initial matching distribution

The M-step updating formula for the initial matching distribution is:

$$m_0(l,k \mid x) = \frac{\sum_{i=1}^{N} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\ \mathbb{1}\Big\{k_{j(i,1)} = k,\ x_{i1} = x\Big\}}{\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{i=1}^{N} p_i(l' \mid \theta^{(m)}, x_{i1}, C)\ \mathbb{1}\Big\{k_{j(i,1)} = k',\ x_{i1} = x\Big\}}.$$

D.2 Wage Distribution Parameters

Wages are log-normal with worker-type and firm-class specific mean and variance.

$$\log f(y_{it} \mid l,x,k) = -\log\sigma_{lxk} - \frac{1}{2}\log(2\pi) - \frac{1}{2}\left(\frac{y_{it}-\mu_{lxk}}{\sigma_{lxk}}\right)^2.$$

The wage segment of the expected log-likelihood writes as follows.

$$W = \sum_{i=1}^{N}\sum_{l=1}^{L} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\sum_{k=1}^{K}\sum_{t=1}^{T_i}\mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x\Big\}\log f(y_{it} \mid l,x,k)\,.$$

Taking derivatives with respect to μ_{lxk} and σ_{lxk}, we obtain the M-step updating formulas for the wage parameters:

$$\mu_{lxk}^{(m+1)} = \frac{\sum_{i=1}^{N} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\sum_{t=1}^{T_i}\mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x\Big\}\, y_{it}}{\sum_{i=1}^{N} p_i(l \mid \theta^{(m)}, x_{i1}, C)\sum_{t=1}^{T_i}\mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x\Big\}}. \tag{D.1}$$

$$\sigma_{lxk}^{(m+1)} = \sqrt{\frac{\sum_{i=1}^{N} p_i(l \mid \theta^{(m)}, x_{i1}, C)\sum_{t=1}^{T_i}\mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x\Big\}\Big(y_{it} - \mu_{lxk}^{(m+1)}\Big)^2}{\sum_{i=1}^{N} p_i(l \mid \theta^{(m)}, x_{i1}, C)\sum_{t=1}^{T_i}\mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x\Big\}}}. \tag{D.2}$$

D.3 Mobility Parameters

Let $M_{lxk\neg}$ denote the probability of staying in k, and $M_{lxkk'}$ the probability of transitioning from k to k', for a worker of type (l, x) and any $k, k' \in \{0, \ldots, K\}$. The mobility component of the expected log-likelihood is:

$$Q = \sum_{i=1}^{N} \sum_{l=1}^{L} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \left[\sum_{t=1}^{T_i-1} \sum_{k=0}^{K} \mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x,\ s_{it} = 0\Big\} \log M_{lxk\neg} \right.$$
$$\left. + \sum_{t=1}^{T_i-1} \sum_{k=0}^{K} \sum_{k'=0}^{K} \mathbb{1}\Big\{k_{j(i,t)} = k,\ k_{j(i,t+1)} = k',\ x_{it} = x,\ s_{it} = 1\Big\} \log M_{lxkk'} \right].$$

Let k denote the firm class at time t, and k' the firm class at time $t+1$. The parametric specification is:

- Unemployment-to-employment transitions: $M_{lx0k'} = \psi_{lxk'}$, for $k' \geq 1$.
- Employment-to-unemployment transitions: $M_{lxk0} = \delta_{lxk}$, for $k \geq 1$.
- Job-to-job transitions: $M_{lxkk'} = \lambda_{lxk'} P_{lxkk'}$, for $k, k' \geq 1$, where $P_{lxkk'} = \dfrac{\gamma_{lxk'}}{\gamma_{lxk'} + h_{lx}\gamma_{lxk}}$.

For $l \in \{1, \ldots, L\}$ and $k, k' \in \{0, 1, \ldots, K\}$, define:

$$n^{(m)}_{lxk\neg} = \sum_{i=1}^{N} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \sum_{t=1}^{T_i-1} \mathbb{1}\Big\{k_{j(i,t)} = k,\ x_{it} = x,\ s_{it} = 0\Big\},$$

$$n^{(m)}_{lxkk'} = \sum_{i=1}^{N} p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \sum_{t=1}^{T_i-1} \mathbb{1}\Big\{k_{j(i,t)} = k,\ k_{j(i,t+1)} = k',\ x_{it} = x,\ s_{it} = 1\Big\}.$$

For a given worker type l, the mobility segment of the expected log-likelihood is:

$$Q = \sum_{k=0}^{K} n^{(m)}_{lxk\neg} \log M_{lxk\neg} + \sum_{k=0}^{K} \sum_{k'=0}^{K} n^{(m)}_{lxkk'} \log M_{lxkk'}.$$

D.3.1 E-NE and JTJ transition probabilities

For a given l, the remaining segment of the expected log-likelihood writes:

$$\sum_{k=1}^{K} n_{lxk\neg}^{(m)} \log(M_{lxk\neg}) + \sum_{k=1}^{K} n_{lxk0}^{(m)} \log(M_{lxk0}) + \sum_{k=1}^{K}\sum_{k'=1}^{K} n_{lxkk'}^{(m)} \log(M_{lxkk'}).$$

Under the parametric specification with inertia, this segment is non-linear in δ, λ, γ, and $h_{lx} = \exp(\alpha_{lx})$. I therefore use a minorising function as in Lentz, Piyapromdee, and Robin (2023), adapted to the inertia term. For $k \in \{1, \ldots, K\}$:

$$M_{lxk\neg} = 1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'} P_{lxkk'} = 1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'} + \sum_{k'=1}^{K} \lambda_{lxk'}\Big(1 - P_{lxkk'}\Big),$$

where

$$P_{lxkk'} = \frac{\gamma_{lxk'}}{\gamma_{lxk'} + h_{lx}\gamma_{lxk}}.$$

A l-type worker stays in class k either because she gets no offer/layoff, or because she gets an offer from k' and rejects it. To build the minorising function, note that at MM iteration s:

$$\begin{aligned} M_{lxk\neg} = {} & \frac{1 - \delta_{lxk}^{(s)} - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s)}}{M_{lxk\neg}^{(s)}} \frac{M_{lxk\neg}^{(s)}}{1 - \delta_{lxk}^{(s)} - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s)}} \left(1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'}\right) \\ & + \sum_{k'=1}^{K} \frac{\lambda_{lxk'}^{(s)}\Big(1 - P_{lxkk'}^{(s)}\Big)}{M_{lxk\neg}^{(s)}} \frac{M_{lxk\neg}^{(s)}}{\lambda_{lxk'}^{(s)}\Big(1 - P_{lxkk'}^{(s)}\Big)} \lambda_{lxk'}\Big(1 - P_{lxkk'}\Big). \end{aligned}$$

Using concavity of the logarithm:

$$\begin{aligned} \log(M_{lxk\neg}) = {} & \log\left(1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'} + \sum_{k'=1}^{K} \lambda_{lxk'}(1 - P_{lxkk'})\right) \\ \geq {} & \frac{1 - \delta_{lxk}^{(s)} - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s)}}{M_{lxk\neg}^{(s)}} \log\left(\frac{1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'}}{1 - \delta_{lxk}^{(s)} - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s)}} M_{lxk\neg}^{(s)}\right) \\ & + \sum_{k'=1}^{K} \frac{\lambda_{lxk'}^{(s)}(1 - P_{lxkk'}^{(s)})}{M_{lxk\neg}^{(s)}} \log\left(\frac{\lambda_{lxk'}(1 - P_{lxkk'})}{\lambda_{lxk'}^{(s)}(1 - P_{lxkk'}^{(s)})} M_{lxk\neg}^{(s)}\right) \equiv \log(\underline{M}_{lxk\neg}). \end{aligned}$$

Equality holds if $\lambda^{(s)}_{lxk'} = \lambda_{lxk'}$, $\delta^{(s)}_{lxk} = \delta_{lxk}$, and $P^{(s)}_{lxkk'} = P_{lxkk'}$. Hence the MM algorithm maximizes:

$$H(M \mid \theta^{(m)}) = \sum_{k=1}^{K} n^{(m)}_{lxk\neg} \log(\underline{M}_{lxk\neg}) + \sum_{k=1}^{K}\sum_{k'=0}^{K} n^{(m)}_{lxkk'} \log(M_{lxkk'}).$$

Given $\theta^{(m)}$, define

$$\tilde{n}^{(s)}_{lxkk'} = n^{(m)}_{lxk\neg} \frac{\lambda^{(s)}_{lxk'}\left(1 - P^{(s)}_{lxkk'}\right)}{M^{(s)}_{lxk\neg}}, \qquad \hat{n}^{(s)}_{lxk} = n^{(m)}_{lxk\neg} \frac{1 - \delta^{(s)}_{lxk} - \sum_{k'=1}^{K} \lambda^{(s)}_{lxk'}}{M^{(s)}_{lxk\neg}}.$$

where $\tilde{n}^{(s)}_{lxkk'}$ is the predicted number of (lx)-type stayers who receive an offer from k' but stay in k, while $\hat{n}^{(s)}_{lxk}$ is the predicted number of (lx)-type stayers who stay because they receive no offer and no layoff.

Plugging $\tilde{n}^{(s)}_{lxkk'}$ and $\hat{n}^{(s)}_{lxk}$ into $H(M \mid \theta^{(m)})$, the segment for γ, conditional on $h^{(s)}_{lx}$, is:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K} \tilde{n}^{(s)}_{lxkk'} \log\left(\frac{h^{(s)}_{lx}\gamma_{lxk}}{\gamma_{lxk'} + h^{(s)}_{lx}\gamma_{lxk}}\right) + \sum_{k=1}^{K}\sum_{k'=1}^{K} n^{(m)}_{lxkk'} \log\left(\frac{\gamma_{lxk'}}{\gamma_{lxk'} + h^{(s)}_{lx}\gamma_{lxk}}\right).$$

With a change of indices:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K} \left(\tilde{n}^{(s)}_{lxkk'} + n^{(m)}_{lxk'k}\right) \log\gamma_{lxk} - \sum_{k=1}^{K}\sum_{k'=1}^{K} \left(\tilde{n}^{(s)}_{lxkk'} + n^{(m)}_{lxkk'}\right) \log\left(\gamma_{lxk'} + h^{(s)}_{lx}\gamma_{lxk}\right).$$

Following Hunter (2004), I apply the tangent minorization to the second term:

$$-\log\left(\gamma_{lxk'} + h^{(s)}_{lx}\gamma_{lxk}\right) \geq 1 - \log\left(\gamma^{(s)}_{lxk'} + h^{(s)}_{lx}\gamma^{(s)}_{lxk}\right) - \frac{\gamma_{lxk'} + h^{(s)}_{lx}\gamma_{lxk}}{\gamma^{(s)}_{lxk'} + h^{(s)}_{lx}\gamma^{(s)}_{lxk}}.$$

With an additional change of indices and simple algebra, this yields the update for γ_{lxk} conditional on $h^{(s)}_{lx}$:

$$\gamma^{(s+1)}_{lxk} = \frac{\sum_{k'=1}^{K} \left(\tilde{n}^{(s)}_{lxkk'} + n^{(m)}_{lxk'k}\right)}{\sum_{k'=1}^{K} \left[\frac{h^{(s)}_{lx}\left(\tilde{n}^{(s)}_{lxkk'} + n^{(m)}_{lxkk'}\right)}{\gamma^{(s)}_{lxk'} + h^{(s)}_{lx}\gamma^{(s)}_{lxk}} + \frac{\tilde{n}^{(s)}_{lxk'k} + n^{(m)}_{lxk'k}}{\gamma^{(s)}_{lxk} + h^{(s)}_{lx}\gamma^{(s)}_{lxk'}}\right]}. \tag{D.3}$$

Given $\gamma^{(s+1)}$, inertia is updated by maximizing the part of the objective that depends on h_{lx}:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'}\,\log(h_{lx}) - \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(n^{(m)}_{lxkk'}+\tilde{n}^{(s)}_{lxkk'}\right)\log\left(\gamma^{(s+1)}_{lxk'}+h_{lx}\gamma^{(s+1)}_{lxk}\right).$$

The first-order condition is:

$$\frac{\partial}{\partial h_{lx}}=0 \iff \frac{1}{h_{lx}}\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'} - \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(n^{(m)}_{lxkk'}+\tilde{n}^{(s)}_{lxkk'}\right)\frac{\gamma^{(s+1)}_{lxk}}{\gamma^{(s+1)}_{lxk'}+h_{lx}\gamma^{(s+1)}_{lxk}}=0.$$

Equivalently:

$$f(h_{lx}) \equiv \sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'} - h_{lx}\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(n^{(m)}_{lxkk'}+\tilde{n}^{(s)}_{lxkk'}\right)\frac{\gamma^{(s+1)}_{lxk}}{\gamma^{(s+1)}_{lxk'}+h_{lx}\gamma^{(s+1)}_{lxk}}=0.$$

Suppose there is positive cross-class mobility mass for group (l,x), so that $\sum_{k=1}^{K}\sum_{k'=1}^{K} n^{(m)}_{lxkk'} > 0$. Then $f(h_{lx})$ is continuous on $(0,\infty)$, with

$$\lim_{h_{lx}\to 0^+} f(h_{lx}) = \sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'} > 0,$$

and

$$\lim_{h_{lx}\to\infty} f(h_{lx}) = \sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'} - \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(n^{(m)}_{lxkk'}+\tilde{n}^{(s)}_{lxkk'}\right) < 0.$$

Hence a positive root exists. Moreover, $f(h_{lx})$ is strictly decreasing on $(0,\infty)$, so this root is unique. In practice, I solve for $h^{(s+1)}_{lx}$ by bisection on $(0,\infty)$: I first construct a bracket $[T_L, T_U]$ such that $f(T_L) > 0$ and $f(T_U) < 0$, starting from a very small T_L and expanding T_U until the sign change appears. I then repeatedly halve the interval, each time keeping the half whose endpoints still have opposite signs, and stop when the first-order condition at the current candidate is sufficiently close to zero. Then

$$\alpha^{(s+1)}_{lx} = \log\left(h^{(s+1)}_{lx}\right).$$

For a given l, the part of the expected log-likelihood used to update $\lambda_{lxk'}$ and δ_{lxk} is:

$$\sum_{k=1}^{K} \hat{n}_{lxk}^{(s)} \log\left(1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'}\right) + \sum_{k=1}^{K}\sum_{k'=1}^{K} \tilde{n}_{lxkk'}^{(s)} \log(\lambda_{lxk'}) + \sum_{k=1}^{K} n_{lxk0}^{(m)} \log(\delta_{lxk}) + \sum_{k=1}^{K}\sum_{k'=1}^{K} n_{lxkk'}^{(m)} \log(\lambda_{lxk'}).$$

Taking derivatives with respect to δ_{lxk}:

$$n_{lxk0}^{(m)}\left[1 - \sum_{k'=1}^{K} \lambda_{lxk'}\right] = \left(n_{lxk0}^{(m)} + \hat{n}_{lxk}^{(s)}\right)\delta_{lxk}. \tag{D.4}$$

Taking derivatives with respect to $\lambda_{lxk'}$:

$$\frac{1}{\lambda_{lxk'}} \sum_{k=1}^{K} \left(\tilde{n}_{lxkk'}^{(s)} + n_{lxkk'}^{(m)}\right) = \sum_{k=1}^{K} \frac{\hat{n}_{lxk}^{(s)}}{1 - \delta_{lxk} - \sum_{k'=1}^{K} \lambda_{lxk'}}. \tag{D.5}$$

Explicit δ_{lxk} from (D.4), plug into (D.5), and update:

$$\lambda_{lxk'}^{(s+1)} = \frac{\sum_{k=1}^{K} \left(\tilde{n}_{lxkk'}^{(s)} + n_{lxkk'}^{(m)}\right)}{\sum_{k=1}^{K} n_{lxk0}^{(m)} + \sum_{k=1}^{K} \hat{n}_{lxk}^{(s)} + \sum_{k=1}^{K}\sum_{k'=1}^{K} \tilde{n}_{lxkk'}^{(s)} + \sum_{k=1}^{K}\sum_{k'=1}^{K} n_{lxkk'}^{(m)}}, \tag{D.6}$$

$$\delta_{lxk}^{(s+1)} = \frac{n_{lxk0}^{(m)}\left(1 - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s+1)}\right)}{n_{lxk0}^{(m)} + \hat{n}_{lxk}^{(s)}}. \tag{D.7}$$

For a given $\theta^{(m)}$, the exact MM updates increase $H(M \mid \theta^{(m)})$ at each iteration s. In implementation, I regularize two steps for numerical stability: I smooth the $\tilde{n}_{lxkk'}^{(s)}$ terms entering the first-order condition for h_{lx} with pseudo-counts equal to 0.5 off the diagonal and 1 on the diagonal, and I impose a lower bound $\gamma_{lxk} \geq 10^{-3}$ before renormalizing $\sum_{k=1}^{K} \gamma_{lxk} = 1$. The implemented routine is therefore a regularized approximation to the exact MM step. I iterate the MM algorithm 200 times and then deliver $\delta^{(m+1)}$, $\gamma^{(m+1)}$, $\lambda^{(m+1)}$, and $\alpha^{(m+1)}$.