Worker Sorting and the Gender Wage Gap

Giulia Vattuone*

May 13, 2024

[Click here for the latest version]

Abstract

Around 20% of the gender wage gap is due to women sorting into firms that pay lower wages. Using French matched employer-employee data, I investigate whether these gender differences in sorting reflect differences in preferences or opportunities. I employ a finite mixture approach *à la* Lentz, Piyapromdee, and Robin (2023) to estimate a model of wages and mobility. Using information on wages, mobility, and observed characteristics, this model classifies workers and firms into a finite number of types and classes. I allow wage profiles and mobility patterns of men and women of the same type to vary over different stages of workers' careers. Counterfactual analyses reveal that over half of the sorting component of the wage gap is driven by differences in preferences, which are more salient among high-wage, mid-experience types. Differences in the job offer distribution following periods of non-employment across all types explain the remaining part.

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 am grateful to Roland Rathelot and Manuel Bagues for their invaluable guidance and infinite patience. For their incredible help and support, special thanks to Christine Braun, Mirko Draca, Emma Duchini, Roberto Pancrazi, and Natalia Zinovyeva. For insightful conversations and feedbacks, I thank Christian Belzil, Luis Candelaria, Julien Combe, Oren Danieli, James Fenske, Rafael Lalive, Jonas Maibom, Christian Moser, Suphanit Piyapromdee, Eric Renault, Seth Sanders, Benoit Schmutz, Isaac Sorkin, Thijs van Rens and participants in seminars at Warwick University, CREST-ENSAE, and Stockholm University. I thank CASD for 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. In this paper, I address this question by estimating a model of wages and mobility by exploiting information on firm-to-firm transitions.

I employ a revealed preference argument in a random search framework, initially proposed by Sorkin (2018). Following this approach, data on observed firm-to-firm transitions are informative about offer arrival rates and worker preferences. The intuition is that, 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 move to a specific firm, while perceived firm values represent worker preferences.

Sorkin (2017) studies revealed preferences through firm-to-firm mobility to estimate gender-specific firm-level values for workers and compares these values to gender-specific firm-level earnings to study the role of compensating differentials in explaining wage inequality between men and women. The novelty of this paper is that it allows for worker heterogeneity within and between genders in a framework that generates rich sorting patterns. I employ a finite mixture model recently proposed by Lentz, Piyapromdee, and Robin (2023) and Bonhomme, Lamadon, and Manresa (2019) and rely on matched employer-employee monthly data for the region *Ile-de-France* (greater Paris) over the period 2015-2019.

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 analyze 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, 2021), 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). 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 like human capital accumulation, psychological attributes or 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.

Administrative data directly provide worker and firm matches, making it challenging to disentangle choices from opportunities to move. There would be one data point for two parameters of interest. The identification of the two key mobility channels requires additional assumptions. First, workers and firms are associated with a finite number of *types* and *classes*, respectively. Second, workers of a given type share the same preferences over firms of a given class, up to an idiosyncratic utility draw specific to the worker-firm match. When choosing between two firms, workers consider the firm's shared value, which is worker-type and firm-class specific, as well as the idiosyncratic utility draw. An interpretation of the idiosyncratic draw is that the choice to move may be influenced by moving costs.

Consider a simplified example to see how these two assumptions disentangle offer arrival rates from preferences. Suppose there is one type of worker and two classes of firms, A and B. Workers, in expectations, are indifferent between firms belonging to the same class. With no loss of generality, we can assume that when workers employed in a firm of class A draw an offer from another firm that also belongs to class A, half of them accept. The expected number of offers from class-A firms is, then, twice as much as the number of transitions that occur *within* class A. We identify the expected number of offers from class B with similar reasoning. Once we pin down the expected number of offers, we can look at the number of *between-class* moves to recover the expected share of workers choosing A over B, and vice versa. Choice probabilities reveal preferences under the following argument: conditional on receiving an offer, if a higher share of workers accepts offers from class A than offers from class B, then we can infer that workers prefer firms in class A. We can extend this simplified example to cases with multiple worker types and firm classes.

Under the aforementioned identifying assumptions, firm-to-firm transition probabilities are modelled as the product of an offer arrival rate and a choice probability. 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.

I estimate the model in two steps as in 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. I allow for flexible interactions between latent types and combinations of gender, tenure, and experience categories. Doing so permits wage profiles and mobility patterns of men and women of the same type to vary over different stages of workers' careers. I follow the iterative process developed by Lentz, Piyapromdee, and Robin (2023) to deal with non-linearities in the mobility parameters. I order workers from low-wage to high-wage types and firms from low-paying to high-paying classes through a standard re-labelling of the estimated clusters.

The clustering results reveal distinct wage gradients based on experience, indicating that latent types capture diverse career trajectories. While wages of low-wage workers are stagnant regardless of experience levels, higher-wage workers earn more with increased experience. The gender wage gap widens with experience among high-wage workers, echoing results from Goldin (2014) and Goldin, Kerr, Olivetti, and Barth (2017). Overall, the residualized gender wage gap is 11 log points, exceeding 30 log points among high-experience, high-wage workers employed in high-paying firms. I find strong correlations between female and male estimated wages calculated over firm classes, implying that firms that pay high wages tend to pay high wages to both men and women.

Using the estimated mobility parameters, I obtain the stationary worker-firm allocations. There are notable differences in where men and women work. In particular, high-experience, high-wage men are more likely than their female counterparts to work at high-paying firms. By simulating a counterfactual scenario where women are allocated across firms as men, I find that gender differences in sorting, the so-called sorting effect, explains 20% of the residualized gender wage gap, consistent with previous studies.

The paper aims to decompose the sorting component of the gender wage gap by quantifying the relative importance of multiple mobility channels. To this purpose, 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*) transitions into non-employment, and *iv*) offer arrival rates while in non-employment. I study how the gender wage gap changes under these multiple scenarios.

First, 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. Higher-experienced workers are likely to drive this effect. 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.[2] 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, gender differences in worker preferences over firms account for over half of the sorting effect. Male workers demonstrate a greater inclination to sort along the wage dimension. A simple correlation analysis suggests that the preference misalignment is most substantial among men and women who are high-wage workers at mid-experience stages, thus likely to be the main contributors to this effect.

Third, differences in transitions into non-employment do not contribute to the gender wage gap. I find strong correlations in the estimates of the exit parameter between female and male workers across all types.

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, especially among those with a strong alignment in the preference parameter. Following periods of

2. (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/

non-employment, women are more likely to draw job offers that pay them less.

Do these sorting effects vary at different career stages? Gender differences in sorting are relatively more important among less experienced workers. They explain 25% of the wage gap among juniors and 16% among seniors with over 20 years of experience. Across all experience groups, the primary determinants are preferences and re-entry rates. A clear preference misalignment is evident among high-wage, mid-experience men and women in their mid-30s. Re-entry patterns between female and male workers are completely misaligned across all worker types, again among those more likely to be in child-rearing ages. These results may thus reflect how the so-called child penalty (Kleven, Landais, and Søgaard, 2019; Adda, Dustmann, and Stevens, 2017) affects differences in mobility that translate into gender wage differentials. Recent evidence shows mothers opt for unemployment insurance benefits and forgo less generous standard parental leave programs (Zurla, 2022). Based on my estimates, patterns in re-entry rates are associated with a wage penalty for women.

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, I complement 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) adopts a revealed-preference approach and concludes that there is an overall agreement between men and women on how they value firms and attributes differences in between-firm pay gaps to differences in the job offer distribution. 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 allow for rich sources of worker heterogeneity. In particular, I allow female and male wages and mobility to vary differently over one's career within a type of worker. Worker types capture different market segments and interact with time-varying covariates in a way that can differ between the genders. 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 .[3]

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, 2022), 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).[4] 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.[5] Throughout the paper, I do not take a stand on whether gender differences in

3. 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.

4. 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.

5. 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

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.

2 Theoretical Framework

This section presents a theoretical framework with which 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. I model firm-to-firm mobility as a function of opportunities to move and preferences. Sorting is intended as the stationary worker-firm allocation and is obtained using the estimates of the mobility parameters.

Second, I am interested in predicting log-wage distributions of workers across firms. With estimates of worker-firm allocations and log-wage distributions, I can document 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 $k_0 = 0$. I estimate firm classes in Section 3.1.

Workers differ in their observed and unobserved characteristics. The set of observed

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.

characteristics consists of experience 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 $l_i \in \{1, \ldots, L\}$ denotes workers' latent heterogeneity.

Gender interact with experience 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, experience, and tenure observed in time period t. Each t refers to a calendar month. I estimate worker types in Section 3.2.

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 $k_{j(i,t)}$. I denote this probability 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 I model job-to-job transitions as the product between a job sampling probability and a choice probability as in Lentz, Piyapromdee, and Robin (2023). 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.

I formally specify all parameters, along with their identification, in Section 2.1.

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 translates into predicting worker-specific average wages across firms and their job mobility in the labour market. The high number of workers and firms makes estimating the parameters of interest burdensome. In addition, and most importantly, separating offer arrival rates from choice probabilities in matched employer-employee data for any worker-firm combination is not possible. The latent-type framework helps overcome these challenges.

Workers and firms are associated with latent *types/classes* that affect earnings and mobility. Worker latent types interact with worker observed characteristics, allowing for a flexible relationship between their observed and unobserved heterogeneity. The interpretation is that, in expectations, workers of a given type earn similar wages and have similar mobility patterns. The latent class captures the heterogeneity of firms that belong to that class.

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.

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 separability between worker and firm heterogeneity. Similar to Lentz, Piyapromdee, and Robin (2023), stayers and movers share the same wage means.[6]

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 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.

6. 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.

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. The model thus assumes that firm-to-firm mobility reveals preferences, allowing for differences in the opportunity to move (represented by the job offer 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. When choosing between the poacher and the incumbent, workers take into account the common value of the firm as well as the idiosyncratic draw. The idiosyncratic utility draw is specific to the worker-firm match and may capture, for example, a mobility cost. 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', workers move to the poacher with a probability that increases in the ratio between the common value of the poacher and the one of the incumbent. 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 it to firm-level earnings to find the role played by compensating differentials in explaining wage inequality. Sorkin (2017) adopts Sorkin (2018)'s 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 labour perspective, 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 labour perspective, 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).

The following subsection presents the likelihood function and describes the specification of the parameters of interest, which formally outline all the model assumptions.

2.1 The theoretical framework in practice

The observed data for worker i consist of sequences of firm identifiers $\Big(j(i,1),\dots,j(i,T)\Big)$, log-hourly wages $\Big(y_{i1},\dots,y_{iT}\Big)$, mobility indicators $\Big(s_{i,1},\dots,s_{i,T-1}\Big)$, gender, and time-varying tenure and experience categories. Interactions between gender and tenure and experience categories are collected in a time-varying vector x_{it}. The latent data consist of the unobserved heterogeneity types $l_i \in \{1,\dots,L\}$ and $k_{j(i,t)} \in \{0,1,\dots,K\}$, for $i \in \{1,\dots,N\}$, $j \in \{0,1,\dots,J\}$, and $t \in \{1,\dots,T\}$.

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:

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

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 $\Pr\left(l_i, k_{j(i,1)} \;\middle|\; x_{i1}\right)$. The worker's observed characteristics consist of interactions between her gender, $g_i \in \{F, M\}$, and combinations of short/long tenure status and experience 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. I divide experience into four groups: 0 to 5 years, 6 to 10 years, 11 to 20 years, 20+ years. The vector of observed characteristics therefore includes $2 \times 2 \times 4$ categories. The initial matching parameter is left completely unrestricted and estimated using simple frequencies. For notational simplicity, from now onwards, I denote the initial matching distribution as $m_0\left(l, k \;\middle|\; x\right)$.

Within a latent class, each firm is equally likely to be selected. I do not explicit the factor that represents firm-specific sampling in the likelihood as, conditional on a firm classification into classes, it gets simplified out in the expectation step of the EM algorithm used to estimate the posterior probability that worker i is of type l, and it is a simple parameter that enters additively the log-likelihood in the maximisation step of the EM algorithm. The uniform-sampling assumption is thus not problematic. As I proceed in two steps, 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 $t \in \{1, \ldots, T-1\}$, 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, $\Pr\left(k_{j(i,t+1)} \;\middle|\; k_{j(i,t)}, l_i, x_{it}\right)$. For notational simplicity, denote the current firm class by k and subsequent firm class by k'. In addition, denote the transition probability by $m\left(k' \;\middle|\; k, l, x\right)$. The worker stays with probability $m\left(\neg \;\middle|\; k, l, x\right) = 1 - \sum_{k'=0}^{K} m\left(k' \;\middle|\; k, l, x\right)$.

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 $k = 1, \ldots, K$ to a firm of class $k' = 1, \ldots, K$ 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 thus specified as the product between a job sampling probability and a choice probability:

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

where $\lambda_{lxk'}$ represents the probability that a worker of type (l, x) receives an offer by a different firm of class k'.[7] Upon receiving an offer, the worker evaluates both the current firm of class k and the potential poacher of class k'. The worker takes into account the values of the firms, common to worker types and firm classes, as well as an idiosyncratic utility draw. The worker moves if the firm of class k' is preferred over the firm of class k. Assuming the idiosyncratic draw is distributed according to a type I extreme value, the choice probability $P_{lx}\Big(k' \succ k\Big)$ is increasing in the ratio of the two common values $\gamma_{lxk'}/\gamma_{lxk}$. The choice probability is therefore an increasing function of the ratio of the perceived common value of the poacher over the perceived common value of the incumbent. To be precise, γ_{lxk} $\forall\ k \in \{1, \ldots, K\}$ is a monotonic transformation of the firm values.

For the estimation, Lentz, Piyapromdee, and Robin (2023) see the choice probability, $P_{lx}\Big(k' \succ k\Big)$, as a Bradley-Terry specification (Bradley and Terry, 1952; Hunter, 2004). 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 the nodes are firms and the edges are non-negative integers of worker transitions between any pair of firms, the Bradley-Terry specification turns 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 obtain a firm ranking that orders firms based on their value. The ranking is obtained for those firms such that

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

there is a path from j to j', for all nodes j and j'. Under the latent-type framework the graph connectivity condition is likely to hold, and it does not require to focus on the set of strongly connected firms. Indeed, I end up having (l, x)-specific $K \times K$ matrices where rows represent arrival firm classes and columns represent departing firm classes. Each cell contains information on the total number of transitions of workers of type l with characteristics x between firm classes. 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 $\sum_{k=1}^{K} \gamma_{lxk} = 1$.

It is worth highlighting that the estimates of γ_{lxk} do not simply represent a ranking of preferences for firm classes. What matters is how much more a firm class is preferred over another. Ratios of firm class values determine how fast workers climb their specific job ladders.

Conditional on a firm classification, it is assumed that in expectations workers are indifferent between two firms belonging to the same class. With no loss of generality, the choice probability between two firms belonging to the same class is assumed to be one half. Under the discretisation of unobserved heterogeneities, the offer arrival rate parameter $\lambda_{lxk'}$ and perceived value γ_{lxk} are identified using information on the frequencies of transition probabilities $m\Big(k' \;\Big|\; k, l, x\Big)$, together with the normalisation $\sum_{k=1}^{K} \gamma_{lxk} = 1$. First, λ_{lxk} is identified for any combination (l, x, k) using data of within-class transitions $m\Big(k \;\Big|\; k, l, x\Big) = \lambda_{lxk} \dfrac{1}{2}$.[8] Second, the choice probabilities $P_{lx}\Big(k' \succ k\Big)$ are pinned down using information from the unrestricted transitions $m\Big(k' \;\Big|\; k, l, x\Big)$ and given knowledge of λ_{lxk}, for any l, x, k. Finally, given the normalisation $\sum_{k=1}^{K} \gamma_{lxk} = 1$, the ratios $\frac{\gamma_{lxk'}}{\gamma_{lxk}}$ follow:

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

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 we may infer that group of workers prefer firms of class k' better than firms of class k. This is the basic principle behind the worker-specific firm values estimation, and this is what is intended by preferences throughout the paper.

8. Under the assumption of no zero cells in the worker-specific job-to-job transition matrices.

Transitions to and from non-employment. At any time period $t \in \{1, \ldots, T-1\}$, a worker of type l and with characteristics x moves from a firm of class $k = 1, \ldots, K$ 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 $k' = 1, \ldots, K$ 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, it follows that the probability of staying into non-employment is:

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

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

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

Hourly wage distributions

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

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

Earnings and hours are recorded at annual frequency, that is there is only one payroll recorded for each employment spell in a year. I calculate hourly wages, annual earnings divided by number of hours, and consider that as the hourly wage in a given month. 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.[9]

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.

9. 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.

I choose $K = 15$, seeking a balance between minimising total intra-class variation and ensuring sufficient observations to fit the data well. Appendix E presents more details about the k-means algorithm and its validation.

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 Figure A1 of Appendix A.[10]

Expectation step

For given parameters $\theta^{(m)}$ and a firm classification C, compute the posterior probability that worker i is of type $l = 1, \ldots, L$:[11]

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

Maximisation step

Maximise the expected log-likelihood with respect to the parameter of interest θ:

$$\sum_i \sum_l p_i\Big(l \;\Big|\; \theta^{(m)}, x_{i1}, C\Big)\; ln\mathcal{L}_i\Big(\theta \;\Big|\; l_i, x_{i1}, C\Big)$$

where k refers to the firm class at t, and k' refers to the firm class at $t+1$, for any $t = 1, \ldots, T$. The maximisation step gives the updated $\theta^{(m+1)}$, used to update the posterior probability in equation 2. Iterate between the expectation step and the maximisation step until convergence.

The maximisation step updating formulas for the wage distributions are simple weighted

10. 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 a few observations in the cells dramatically reduce the estimation's precision. Figure A1 shows that the selected number of types fits the data well.

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

averages and variances using the posterior probabilities as weights. Simple frequencies are the maximisation step for the initial matching distribution and unrestricted transition probabilities. Transition probabilities for job-to-job mobility are non-linear in the parameters of interest. In Appendix F, I detail the minorisation-maximisation (MM) algorithm proposed by Lentz, Piyapromdee, and Robin (2023) to maximise the expected log-likelihood for the non-linear cases.

4 Data and Sample Selection

I use the French matched employer-employee data, *Déclarations Annuelles de Données Sociales* (DADS), over the period 2015-2019. These datasets are collections of mandatory employer reports of salaried employees, compiled by the French statistical institute *Institut National de la Statistique et des Etudes Economiques* (INSEE). The data contain job-spell-level information on worker-firm matches. Importantly, working hours are reported, allowing me to control for gender differences in labour supply. In order to estimate the model described in Section 2, I use two data sources: the DADS-Postes and the DADS-Panel.

DADS-Postes

The DADS-Postes dataset contains information on the universe of jobs in France. It lacks a proper longitudinal dimension, as the worker identifiers change every two years. Nonetheless, it is helpful for the firm clustering described in section 3.1, as it provides complete yearly employment information for each firm.

To select firms for the analysis, I focus on those that meet the following criteria: they employ at least one full-time worker in Ile de France in 2015, employ both genders, and have been active for five consecutive years. I consider as employment any job spell with positive wages and hours. Wages are reported at an annual frequency.[12] Table B1 compares firms across different selection steps. Column 1 describes firms that employed at least a full-time job in Ile de France in 2015, and Column 2 describes those hiring both genders and being active for five consecutive years. The restrictions in Column 2 result in focusing on

12. I winsorize hourly wages at 0.001 and 0.999

firms that, on average, have twice the number of workers, pay higher wages, require longer hours, and where workers have higher chances of holding managerial positions. Column 3 describes firms where I observe workers in the DADS Panel. As a result of this additional restriction, the firms in the final sample have, on average, 250 employees and pay 40% higher wages compared to the full sample of firms that employ at least a full-time job in Ile de France in 2015. They have slightly a lower gender balance but they are 66% more likely to have women in managerial positions. The final dataset includes information on firms' gender composition, wage distributions, workers' basic demographics, occupations, sectors, and public/private status. In section 5, I present ex-post tabulations of firms' observed characteristics by predicted latent class.

DADS-Panel

The DADS-Panel dataset contains information on the employment history of workers born in October. I use the DADS-Panel for the maximum likelihood estimation conditional on firm clustering.

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, worker experience in the labour market, starting and ending day of the spell in the year. I construct a monthly panel of individual working trajectories, considering job spells with positive wages and hours. I obtain hourly wages by dividing yearly earnings by the hours worked. I select workers employed in January 2015, only working in *Ile-de-France* over the period 2015-2019, who have never worked in the Agricultural sector, who hold 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 firms selected from the DADS-Postes dataset over the period 2015-2019. I track the selected workers over time. Non-employment is the time between different job spells.[13] Table B2 compares workers across different selection steps. The restrictions I introduce select the sample in terms of higher average earnings, longer working hours, and the likelihood of holding managerial positions.

13. 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 minimise the probability that non-employment gets confounded with periods of education.

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.

Table 1 presents the selected sample, consisting of 49% women and 51% men. Women, on average, earn 23% less than men and work 5% fewer hours, with a higher prevalence of part-time employment. While both genders display similar probabilities of monthly job transitions, women exhibit a slightly higher propensity to transition from employment to non-employment and are more likely to transition from non-employment to employment. Regarding occupational positions, 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 monthly length of job spells spans three and a half years for both genders.

Table 2 provides a detailed breakdown by gender, experience, and tenure. Earnings tend to increase with experience and tenure, although more so for men than women. Among short-tenured workers with less than five years of experience, women earn €6,500 less than men, a gap that widens to €13,600 for long-tenured workers with over 20 years of experience. The disparity in working hours also widens with experience, particularly among short-tenured workers. The proportion of women engaged in part-time employment surpasses that of men across all experience and tenure categories, with this gap accentuating as experience grows. Finally, while mobility rates between firms and average job spell durations remain similar between genders, women exhibit higher rates of transitioning into and out of employment across all experience and tenure groups.

5 Firm Classes and Worker Types

This section presents ex-post tabulations of firm and worker observed heterogeneity by predicted latent firm classes k and worker types l, respectively. To facilitate the interpretation, firm classes and worker types are relabelled so that they are increasing in the estimated log-hourly wage $\hat{\mu}_{lkx}$.[14] I therefore refer to higher latent groups as higher-paying firms and higher-wage workers.

Firm classes

Table 3 provides descriptive statistics of the average firm within each cluster throughout the entire sample period from 2015 to 2019. Firm classes exhibit significant disparities in size and gender composition, with some classes predominantly employing women while others lean towards hiring men. Despite these variations, no clear correlations emerge between gender compositions and metrics such as wages, working hours, or industry sectors. A pattern emerges regarding career progression: in classes where women comprise the majority of employees, women tend to exhibit higher probabilities of assuming managerial roles than men and constitute the majority among managers.

There is a clear trend wherein high-paying classes offer higher wages to both men and women and longer working hours. Finally, firms operating in sectors like construction and finance are more likely to belong to high-paying classes, while those in hospitality, education, and health sectors tend towards lower-paying classes.

Worker types

Table 4 describes workers' wages and mobility rates by predicted worker types. It offers insights into how latent types interact with observed characteristics, presenting a detailed breakdown by experience, tenure, and gender categories. Upon comparing the different worker types, distinct experience gradients stand out: low-wage workers exhibit stagnant wages irrespective of experience levels, whereas mid-wage and high-wage workers have higher salaries with increased experience. Gender differentials in wages and hours are evident, with women consistently earning and working less than their male counterparts. The wage gap widens with experience, from 15 to 24 log points among short-tenured high-wage workers

14. Latent worker types l and latent firm classes k are per se meaningless, they simply capture unobserved heterogeneity of groups of workers and firms. I consider the standard two-way fixed effects projection $Y = \nu_x + \alpha_l + \psi_k + \epsilon$ run on a cross-sectional database obtained from the empirical matching distribution $\Pr\big(l, x, k\big)$ (see Section 6). Where ν_x are interactions among observable characteristics, α_l is the worker effect and ψ_k is the firm effect. I relabel l and k so that α_l and ψ_k are increasing in l and k, respectively. The relabelling allows to interpret higher l as higher-wage worker types, and higher k as higher-paying firm classes.

and from 19 to 28 log points among long-tenured counterparts, underscoring the expanding nature of the gap over the life cycle (Goldin, 2014; Goldin, Kerr, Olivetti, and Barth, 2017).

The table indicates that latent types capture diverse career trajectories among workers, highlighting gender divergences over high-wage type careers. Regarding mobility, workers within a latent type tend to relocate less as experience accumulates, but high-wage workers exhibit higher transition rates than lower-wage workers within specific experience groups. Finally, women are equally represented across all categories, as also documented in Table C1.

Gender differences in wages and worker-firm allocations

Figure 1 compares the estimated log wages between female and male workers across various worker types, firm classes, and experience-tenure categories. The colour gradient denotes the progression from low-wage workers in low-paying firm classes (light grey dots), to high-wage workers in high-paying firm classes (black dots). The different shapes of dots distinguish between various experience categories, which include both short and long tenures. The arrangement of dots reveals a strong correlation between female and male wages, formally calculated in Panel A of Table 5. The distance between estimate points and the 45-degree dashed line represents the gender wage gap. Despite a few exceptions, men systematically earn higher salaries, especially among high-experience, high-wage workers in high-paying firms where the gap exceeds 30 log points.

Regarding the allocation of female and male workers across firm classes, two key insights emerge from Figure 2. First, men and women are unequally distributed across firms. High-experience, high-wage men are more likely than their female counterparts to work at high-paying firms. Second, there is no solid evidence of strong wage sorting for either gender, indicating that high-experience, high-wage workers do not consistently gravitate towards higher-paying firm classes. However, a closer look may reveal a stronger tendency for wage sorting among men.

In the following Section, I formally describe how I obtain the stationary worker-firm allocations, investigate the primary mobility channels contributing to gender disparities in worker-firm allocations, and conduct counterfactual analyses to quantify their respective contributions to the overall gender wage gap.

6 Worker Sorting and the Gender Wage Gap

In this section, I explore gender-based differences in worker sorting across firms. The aim is to evaluate the *sorting effect*, which measures the portion of the gender wage gap attributed to the unequal distribution of men and women among firms. In addition, the ultimate goal is to quantify the relative importance of patterns in offer arrival rates, exit rates, and worker preferences over firms in contributing to the sorting effect. To achieve this, I derive the stationary matching distribution from the predicted worker-firm stationary allocations and worker-type marginal distributions. Subsequently, I run a series of counterfactual exercises to simulate scenarios where female workers move across firms at the same rates of male workers.

The worker-firm stationary allocation

Sorting is the stationary allocation of worker types and firm classes, $\Pr^*\Big(k \;\Big|\; l, x\Big)$, for any $l \in \{1, \dots, L\}$, $k \in \{0, 1, \dots, K\}$, and combinations of observed characteristics x which include interactions of short/long tenure, experience and gender categories. For each (l, x) combination, I build a $(K+1) \times (K+1)$ transition matrix, M_{lx}, with k'-th row and k-th column cell corresponding to the (l, x)-specific estimated probability of moving from $k = 0, 1, \dots, K$ to $k' = 0, 1, \dots, K$. The stationary allocation is the eigenvector corresponding to an eigenvalue of one of the transition matrix M_{lx}.

The stationary matching distribution

I obtain the empirical distribution of matches, $\Pr\Big(l, x, k\Big)$, using the sorting distribution and worker type frequencies:

$$\Pr\Big(l, x, k\Big) = \Pr^*\Big(k \;\Big|\; l,\ x\Big) \sum_{i=1}^{N} p_i\Big(l \;\Big|\; \theta^{(m)}, x_{i1}, C\Big) \sum_{t} \mathbb{1}\Big\{x_{it} = x\Big\}$$

where $\Pr^*\Big(k \;\Big|\; l, x\Big)$ is the stationary allocation of worker types and firm classes and the second element of the multiplication, under a normalisation, is just $\Pr\Big(l, x\Big)$. I augment the initial sample size by L and simulate a cross-sectional dataset from the empirical distribution of matches $\Pr\Big(l, x, k\Big)$, in which log-hourly wages are drawn using the estimated log-normal distributions centred in $\hat{\mu}_{lxk}$ with variance $\hat{\sigma}^2_{lxk}$. Workers earn no wage in non-employment.

The sorting effect and its determinants

The *sorting effect* is the percentage change difference between the gender wage gap estimated with gender-based differences in worker-firm allocations and the gender wage gap estimated under a counterfactual scenario in which men and women are equally distributed across firms. I obtain the counterfactual distribution of matches by equating the worker-firm allocations of women to the ones of men. The counterfactual distribution of matches is:

$$\tilde{\Pr}\Big(l, x, k\Big) = \tilde{\Pr}^*\Big(k \;\Big|\; l, x\Big) \Pr\Big(l, x\Big)$$

I obtain $\tilde{\Pr}^*\Big(k \Big| l, x\Big)$ by equating the mobility patterns of women to the ones of the corresponding (l, x) male category. In practice, I simulate two cross-sectional datasets: one from $\Pr\Big(l, x, k\Big)$, one from $\tilde{\Pr}\Big(l, x, k\Big)$. The percentage change difference between the gender wage gaps estimated from the two datasets is the *sorting effect*.

I quantify the relative importance of offer arrival rates while employed, exit rates, offer arrival rates while non-employed, and workers' perceived firm values by equating the estimated mobility parameters of female workers to the ones of male workers. Whenever I equate a mobility parameter, I predict the stationary worker-firm allocations, obtain the corresponding empirical matching distribution, simulate a cross-sectional dataset, draw wages, and compute the gender wage gap.

Results

Figure 3 presents five counterfactual exercises run on the full cross-sectional dataset. Table D1 reports the corresponding point estimates and standard errors. I describe the results reading from left to right.

First, when considering the estimated worker-firm allocations, where men and women are distributed unequally across firms, the residualized gender wage gap stands at 11 log points.

Second, if employed female workers faced the job offer distribution of their male counterparts, the gender wage gap would increase to 12.4 log points. In light of the strong correlation between wages of both genders over firm classes, shown in Panel A of Table 5, this result suggests that women may be more likely to receive offers for higher-paying positions. The orthogonal nature of job offer rates between female and male high-experience, high-wage workers, as documented in Panel B of Table 5, further supports this interpretation. 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.[15] To the extent that callback rates reflect actual job offers and executive roles are more likely to be held by high-experienced workers, the results presented in Figure 3 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.

Third, conditional on facing the male job offer distribution, if women had preferences over firms similar to men, accepting offers at the same pace, the gender wage gap would diminish to 9.8 log points. This result suggests that gender disparities in firm-to-firm transitions contribute to 11% of the residualized gender wage gap, with differences in preferences being the primary determinant. Panel C of Table 5 examines the correlation between female and male preferences over firms. Interestingly, preferences tend to align strongly within latent types at higher experience levels. However, the correlations are negligible or negative for high-wage workers with 6-10 and 11-20 years of experience. These worker groups are thus more likely to drive the counterfactual result. Female workers in these experience groups are often in child-rearing stages. Through firm-to-firm transitions, they may be willing to pay a price to achieve a better work-life balance to manage family or caregiving responsibilities typically shouldered by women.

Fourth, if women had the same firm-to-firm transition patterns as men, experiencing similar patterns also in exit rates would have virtually no impact on the gender wage gap. Panel D of Table 5 underscores the strong correlations in the exit parameter between male and female workers. Gender disparities in transitions from employment to non-employment do not significantly influence the gender wage gap.

Finally, the fifth counterfactual exercise additionally equates offer rates for non-employed

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

women to those of men. This scenario reflects a situation where women are distributed across firms similarly to men.[16] Under this counterfactual, the gender wage gap would decrease to 8.8 log points. This finding suggests a sorting effect that explains 20% of the residualized gender wage gap, consistent with previous studies on several countries. The two primary determinants of this effect are differences in acceptance rates driven by preferences and differences in offer rates in non-employment, which are equally important. While some aspects of the preference mechanism will almost certainly contribute to re-entry patterns, Panel E of Table 5, in contrast to Panel C, reveals weak correlations between female and male entry rates across *all* worker types, particularly among workers more likely to be in child-rearing ages. The offer distribution that women face while in non-employment is associated with a wage penalty.

Figure 4 replicates the exercises for different first-period experience groups. While initially confirming the overall results across experience levels, three insights emerge. First, gender differences in offer arrival rates favouring women are apparent for higher-experience workers, supporting the conclusions of the second counterfactual regarding senior roles. Second, the sorting effect progressively decreases as experience increases, declining from 25% for young workers with 0-5 years of experience to 16% for senior workers with 20+ years of experience. Third, the relative importance of the preference channel slightly increases with experience.

7 Discussion

This article integrates administrative data with a flexible model to decompose the sorting component of the gender wage gap first presented by Card, Cardoso, and Kline (2016). In particular, it documents the relative importance of multiple mobility patterns and channels in determining gender wage differentials. It examines three types of transitions: firm-to-firm movements, transitions into non-employment, and transitions into employment.

Leveraging the finite mixture approach of Lentz, Piyapromdee, and Robin (2023), the analysis incorporates rich sources of two-sided heterogeneity. Transition probabilities between firms are modelled as the product of offer arrival rates and choice probabilities. Using a

16. Graphically, under this scenario, all dots in Figure 2 would lie on the 45-degree line.

revealed preference argument as in Sorkin (2018), worker preferences guide job offer acceptance decisions: upon receiving a job offer, the higher the perceived value of the poacher, the higher the probability the worker chooses to accept the offer. Offer arrival rates and worker preferences are separately identified under a discretization of worker and firm heterogeneity and a distributional assumption on idiosyncratic factors affecting workers' mobility decisions. The model estimation follows a two-step procedure, similar to Bonhomme, Lamadon, and Manresa (2019).

The study uses French monthly matched employer-employee data and focuses on the Paris region from 2015 to 2019. Counterfactual exercises reveal that if women were distributed across firms as men are, the gender wage gap would reduce by 20 percent. Differences in preferences account for over half of this sorting component of the gender wage gap, while differences in offer rates in non-employment explain the remaining part.

Some caveats are worth discussing. First, the framework does not consider idiosyncratic shocks to wages and layoff notifications as explanations of moves observed in the data, potentially confounding voluntary and involuntary transitions in the revealed preference argument. However, guessing how these omissions lead to biased estimates is not straightforward. The analysis leverages all the flows made by groups of workers across groups of firms to average out idiosyncrasies and extract a 'systematic' pattern of preferences.

Second, the focus is not on long panels that would allow to consider age dynamics as done, for example, in Barth, Kerr, and Olivetti (2021) and Amano-Patiño, Baron, and Xiao (2021). Yet, the worker clustering of this article proves instrumental in allowing for both within and between gender variation. Interacting gender with experience and tenure categories is a way to allow for gender differences in wage profiles and mobility patterns over different career stages within a worker type.

Finally, while sorting based on preferences explains more than 10 percent of the residualized gender wage gap, it is important not to interpret these findings 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, these choices blur what we define as preferences (Bertrand, 2020). The findings of this paper recall proposals entailing restructuring 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`.

Adda, Jérôme, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy* 125 (2): 293–337. 10.1086/690952.

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`.

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, 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.

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. 2022. "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.

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.

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.

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.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209. 10.1257/app.20180010.

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 Guido, Alexandra Roulet, and Mark Stabile. 2021. "Understanding the role of firms in the gender wage gap over time, over the life cycle, and across worker types." Discussion Paper 16671, CEPR.

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.

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.

Zurla, Valeria. 2022. "How Should We Design Parental Leave Policies? Evidence from Two Reforms in Italy." Working Paper.

8 Tables

Table 1: Sample description

Gender:	Women	Men
Mean Annual Earnings	40,874	53,178
Mean Hours	1,693	1,784
Share Part-time	16%	5%
Mean Age	41	41
Share doing JTJ	15%	15%
Share doing E-NE	19%	16%
Share doing NE-E	45%	38%
Share Managers	34%	42%
Share Intermediate	28%	20%
Share Employee non-manual	31%	17%
Share Employee manual	6%	21%
Mean job spell months	43	44
N workers	80,967	84,191

Notes: The table presents descriptive statistics for the selected sample. Data relative to mobility represents the share of workers doing at least one transition of a given type over the sample period. JTJ stands for Job-To-Job. E stands for Employment. NE stands for Non-Employment. The numbers in the table represent averages computed over the pooled period of January 2015 - December 2019.

Table 2: Sample description by gender, experience, and tenure

	Women							
Tenure:	Short				Long			
Years of experience:	0-5	6-10	11-20	20+	0-5	6-10	11-20	20+
Mean Annual Earnings	26,692	30,161	34,122	37,358	32,900	34,721	40,174	44,925
Mean Hours	1,438	1,494	1,503	1,520	1,622	1,659	1,692	1,732
Share Part-time	14%	12%	15%	18%	12%	12%	18%	16%
Mean Age	29	31	38	48	29	31	38	48
Share doing JTJ	11%	9%	8%	8%	10%	13%	10%	7%
Share doing E-NE	14%	13%	12%	11%	12%	17%	12%	8%
Share doing NE-E	36%	32%	27%	24%	32%	31%	30%	20%
Mean job spell months	10	11	12	12	12	23	32	40
N workers	4,774	10,181	14,257	7,609	4,754	20,362	47,476	39,714
	Men							
Tenure:	Short				Long			
Years of experience:	0-5	6-10	11-20	20+	0-5	6-10	11-20	20+
Mean Annual Earnings	33,229	36,829	44,002	48,976	39,212	43,273	52,001	58,543
Mean Hours	1,510	1,596	1,628	1,648	1,695	1,747	1,790	1,810
Share Part-time	9%	6%	5%	6%	6%	5%	4%	4%
Mean Age	29	31	38	48	29	31	38	48
Share doing JTJ	9%	8%	8%	7%	11%	14%	10%	7%
Share doing E-NE	11%	10%	11%	10%	11%	13%	10%	8%
Share doing NE-E	33%	27%	25%	24%	33%	22%	20%	19%
Mean job spell months	11	11	12	12	13	23	32	41
N workers	4,135	8,872	13,711	9,011	4,310	18,640	48,211	44,958

Notes: The table presents descriptive statistics for the selected sample by combinations of gender, experience, and tenure. The sample period is January 2015 - December 2019. Mobility data present the share of workers engaging in at least one transition of a given type during the sample period. *JTJ* stands for Job-To-Job. *E* stands for Employment. *NE* stands for Non-Employment. Earnings and hours data present yearly averages. *N workers* refers to the number of workers who happened to have a given combination of gender, experience, and tenure from January 2015 to December 2019. That is, the same worker can appear in multiple columns.

Table 3: Firm classes description

Firm class:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N firms	570	1,973	1,763	1,576	1,001	1,992	1,121	1,360	2,298	1,017	2,772	1,480	1,939	2,799	2,264
Mean size	294	198	142	248	607	179	818	299	175	460	211	237	197	213	110
Female share	75%	72%	32%	28%	67%	24%	31%	80%	57%	55%	22%	69%	54%	27%	41%
Mean log hourly wage	2.55	2.75	2.69	2.74	2.65	2.91	2.78	2.81	2.94	2.88	3.01	3.13	3.12	3.19	3.50
Mean hours	1,159	1,246	1,232	1,294	1,223	1,383	1,347	1,277	1,279	1,359	1,448	1,409	1,324	1,460	1,495
% Women managers	11%	25%	16%	9%	6%	13%	6%	12%	29%	11%	10%	21%	28%	14%	21%
% Men managers	7%	16%	29%	18%	4%	28%	13%	7%	26%	12%	25%	17%	28%	29%	32%
% Women among managers	56%	62%	35%	30%	55%	30%	32%	64%	53%	47%	27%	54%	49%	31%	38%
% Firms in Hotel	9%	3%	2%	8%	17%	2%	14%	4%	1%	15%	4%	2%	1%	1%	1%
% Firms in Admin Services	13%	6%	9%	15%	13%	6%	13%	6%	6%	6%	6%	6%	7%	4%	5%
% Firms in Construction	0%	0%	3%	2%	0%	13%	4%	0%	1%	1%	17%	2%	2%	17%	9%
% Firms in Commerce	21%	12%	12%	18%	15%	16%	19%	13%	9%	16%	21%	18%	13%	24%	28%
% Firms in Education	8%	30%	3%	4%	4%	1%	2%	7%	22%	4%	1%	7%	11%	4%	2%
% Firms in Managing	9%	11%	19%	9%	2%	12%	4%	5%	15%	6%	7%	15%	20%	10%	13%
% Firms in Finance	2%	3%	2%	1%	0%	2%	1%	2%	6%	3%	2%	8%	12%	5%	11%
% Firms in Pub Admin	2%	5%	0%	1%	28%	0%	3%	15%	2%	11%	1%	8%	1%	2%	2%
% Firms in Health Accomm	10%	8%	2%	1%	10%	0%	2%	22%	2%	5%	0%	3%	1%	0%	1%

Notes: The table presents ex-post tabulations of observed firm characteristics by predicted latent classes. Firm classes are obtained using the k-means algorithm described in Section 3.1, with higher classes corresponding to higher-paying ones. All numbers in the table are obtained considering the entire sample period 2015-2019.

Table 4: Worker types description

Tenure:	Short						Long					
Worker type:	Low-wage		Mid-wage		High-wage		Low-wage		Mid-wage		High-wage	
	F	M	F	M	F	M	F	M	F	M	F	M
Unconditional log wage												
Experience 0-5	2.78	2.87	2.83	2.95	2.88	3.03	2.84	2.94	2.91	3.06	2.97	3.16
Experience 6-10	2.79	2.87	2.91	3.04	2.98	3.18	2.85	2.96	2.95	3.12	3.09	3.33
Experience 11-20	2.78	2.86	2.99	3.15	3.18	3.44	2.88	2.97	3.05	3.21	3.29	3.58
Experience 20+	2.74	2.86	3.04	3.19	3.30	3.54	2.88	2.98	3.11	3.25	3.43	3.71
Hours												
Experience 0-5	1,573	1,638	1,545	1,626	1,530	1,578	1,679	1,715	1,653	1,715	1,603	1,652
Experience 6-10	1,577	1,671	1,578	1,686	1,532	1,641	1,680	1,755	1,678	1,761	1,619	1,721
Experience 11-20	1,554	1,692	1,618	1,742	1,569	1,714	1,692	1,784	1,715	1,806	1,662	1,777
Experience 20+	1,523	1,703	1,654	1,769	1,617	1,746	1,718	1,800	1,758	1,823	1,705	1,792
Share doing JTJ												
Experience 0-5	0.14	0.09	0.08	0.07	0.10	0.09	0.11	0.11	0.08	0.10	0.10	0.11
Experience 6-10	0.10	0.08	0.08	0.08	0.10	0.09	0.13	0.13	0.11	0.12	0.15	0.16
Experience 11-20	0.08	0.07	0.07	0.07	0.09	0.09	0.09	0.09	0.08	0.08	0.13	0.15
Experience 20+	0.08	0.07	0.08	0.08	0.07	0.07	0.06	0.06	0.06	0.06	0.10	0.10
Share doing E-NE												
Experience 0-5	0.12	0.10	0.15	0.10	0.15	0.14	0.10	0.10	0.12	0.11	0.16	0.16
Experience 6-10	0.11	0.09	0.14	0.10	0.16	0.13	0.15	0.11	0.16	0.13	0.23	0.18
Experience 11-20	0.12	0.11	0.11	0.09	0.13	0.11	0.12	0.09	0.09	0.08	0.18	0.15
Experience 20+	0.14	0.11	0.10	0.10	0.10	0.10	0.08	0.08	0.05	0.05	0.13	0.12
Share doing NE-E												
Experience 0-5	0.35	0.37	0.31	0.26	0.41	0.34	0.37	0.35	0.26	0.22	0.32	0.41
Experience 6-10	0.32	0.28	0.27	0.25	0.34	0.27	0.36	0.23	0.26	0.18	0.30	0.23
Experience 11-20	0.27	0.27	0.25	0.23	0.28	0.25	0.33	0.21	0.25	0.15	0.31	0.22
Experience 20+	0.26	0.25	0.23	0.25	0.23	0.23	0.22	0.22	0.15	0.14	0.22	0.19
N workers in $t = 1$	8,493	9,012	6,105	6,689	5,858	5,656	22,043	23,726	21,795	22,964	16,674	16,145
N observations	192,288	191,668	132,597	137,000	160,187	142,313	1,639,905	1,772,588	1,541,349	1,642,131	1,191,694	1,165,760

Notes: The table presents ex-post tabulations of workers' wages and mobility rates by predicted latent types and observed experience, tenure, and gender categories. Workers are classified into types using the EM algorithm described in Section 3.2. Data relative to log wages correspond to the sample period January 2015 - December 2019 and represent averages weighted by the estimated posterior probability $p_i\left(l \middle| \theta^{(m)}, x_{i1}, C\right)$. 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 $\sum_{i=1}^{N} p_i\left(l \,\middle|\, \theta^{(m)}, x_{i1}, C\right) \sum_t \mathbb{1}\left\{x_{it} = x\right\}$.

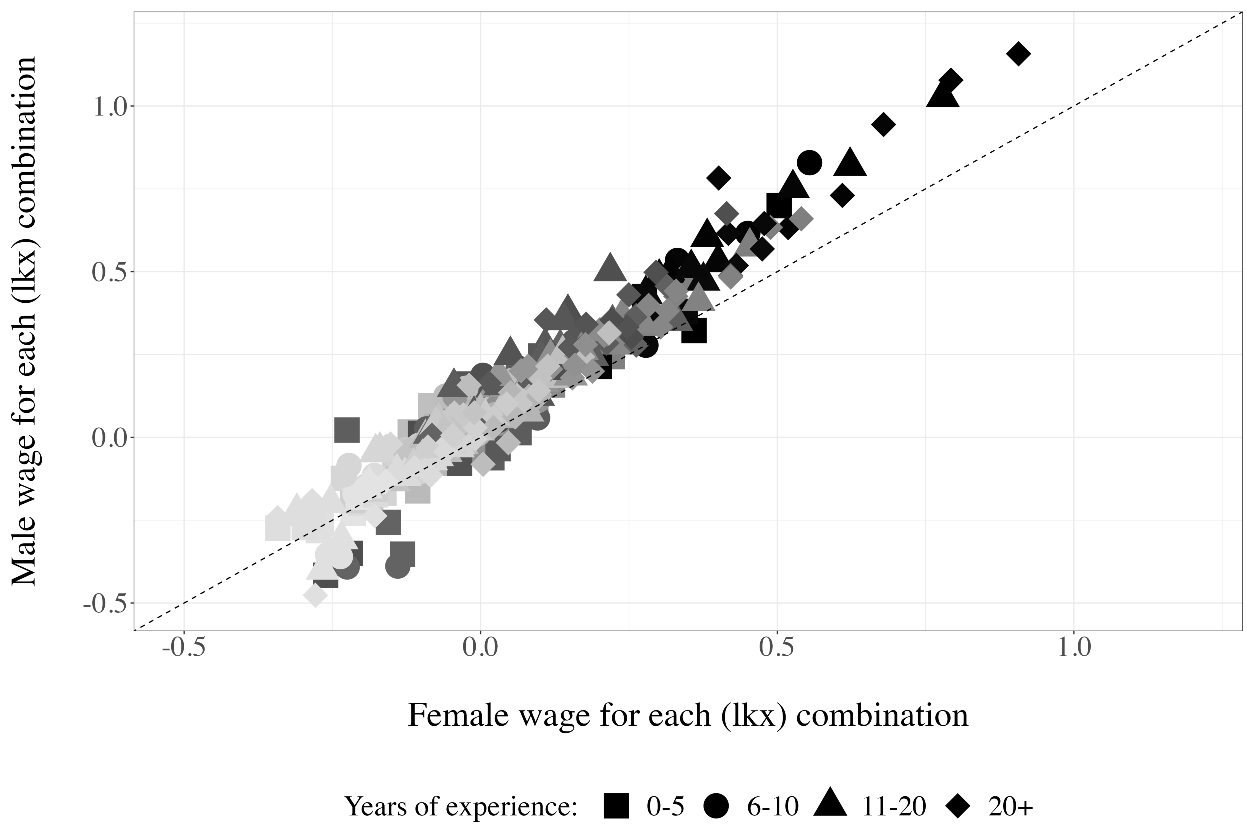
Table 5: Correlations between estimated parameters

Worker type:	Low-wage	Mid-wage	High-wage
Panel A: *Log wage* μ_{lkx}			
Experience 0-5	0.92	0.92	0.90
Experience 6-10	0.89	0.94	0.94
Experience 11-20	0.93	0.96	0.96
Experience 20+	0.92	0.95	0.96
Panel B: *Job offer arrival rate* λ_{lkx}			
Experience 0-5	0.26	0.84	0.54
Experience 6-10	0.60	0.81	0.42
Experience 11-20	-0.09	0.35	0.28
Experience 20+	0.22	0.39	0.03
Panel C: *Preferences* γ_{lkx}			
Experience 0-5	0.41	0.02	0.12
Experience 6-10	0.68	0.30	0.05
Experience 11-20	0.80	0.16	-0.18
Experience 20+	0.77	0.62	0.32
Panel D: *Exit rates* δ_{lkx}			
Experience 0-5	0.12	0.42	0.16
Experience 6-10	0.74	0.73	0.76
Experience 11-20	0.84	0.71	0.78
Experience 20+	0.77	0.76	0.39
Panel E: *Entry rates* ψ_{lkx}			
Experience 0-5	0.24	-0.03	0.43
Experience 6-10	0.01	-0.25	0.10
Experience 11-20	-0.17	-0.01	0.01
Experience 20+	0.20	0.03	0.15

Notes: The table presents correlations between the estimated mobility and wage parameters for women and those for men. The correlations are calculated over firm classes k for given worker types l and experience and tenure categories x, then averaged over short and long tenure.

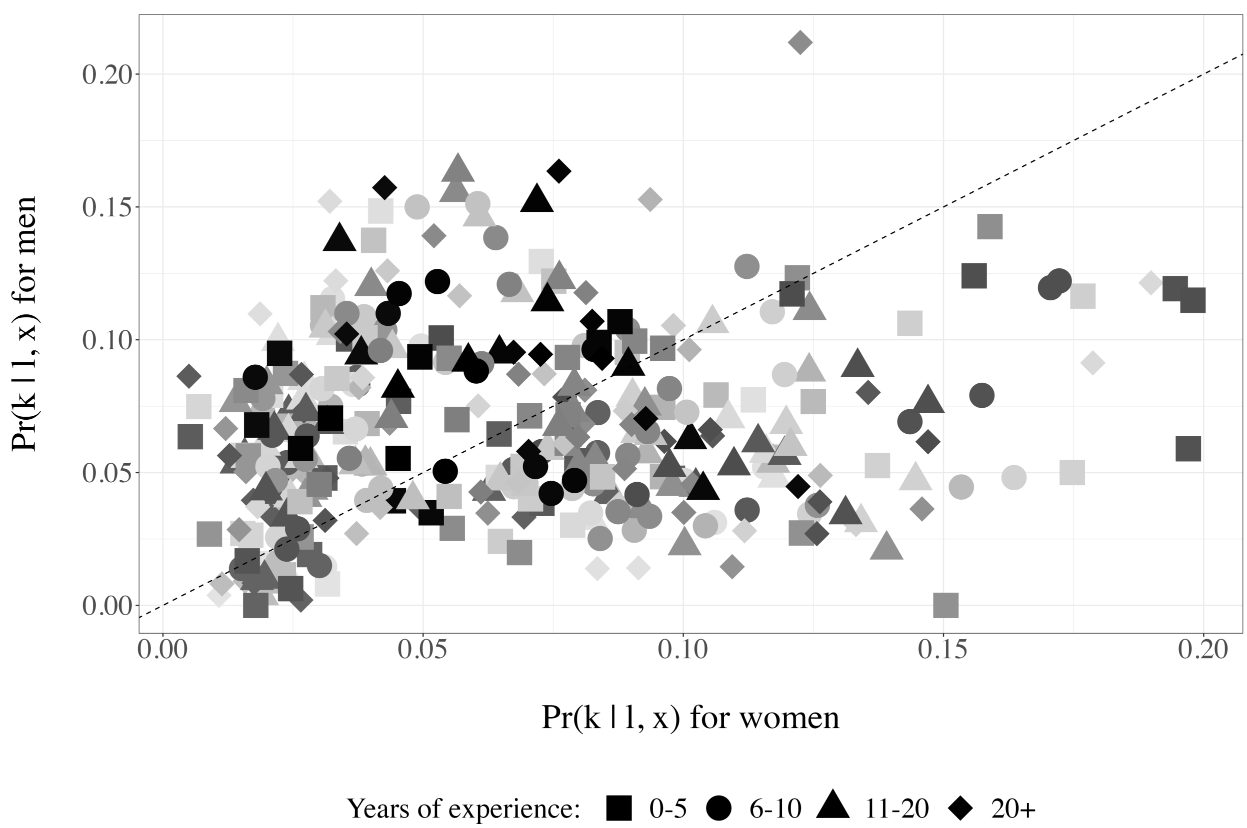
9 Figures

Figure 1: Gender wage differentials



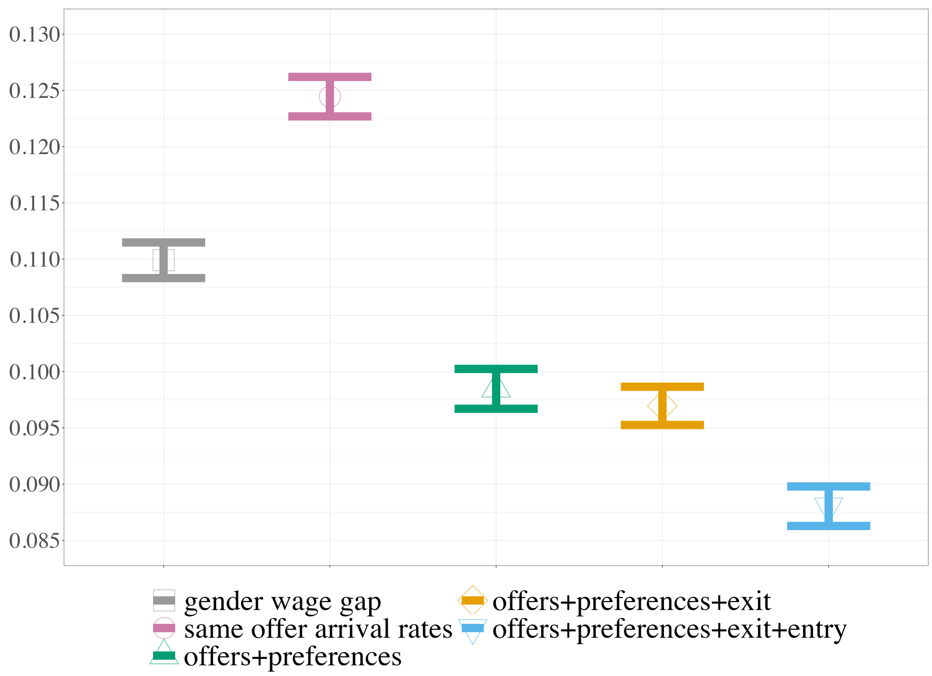
Notes: The figure shows a scatter plot to compare the estimated wages of men and women, $\hat{\mu}_{lkx}$. The x-axis represents the estimated average wage for women, while the y-axis corresponds to the average wage for men. Each dot refers to workers of type l with experience and tenure categories x and employed in firm class k. The shapes on the graph indicate different levels of experience, with the darkness of colour intensifying to indicate higher-wage types and higher-paying firm classes. The dashed line represents the 45 degree line.

Figure 2: Gender differences in worker-firm allocations



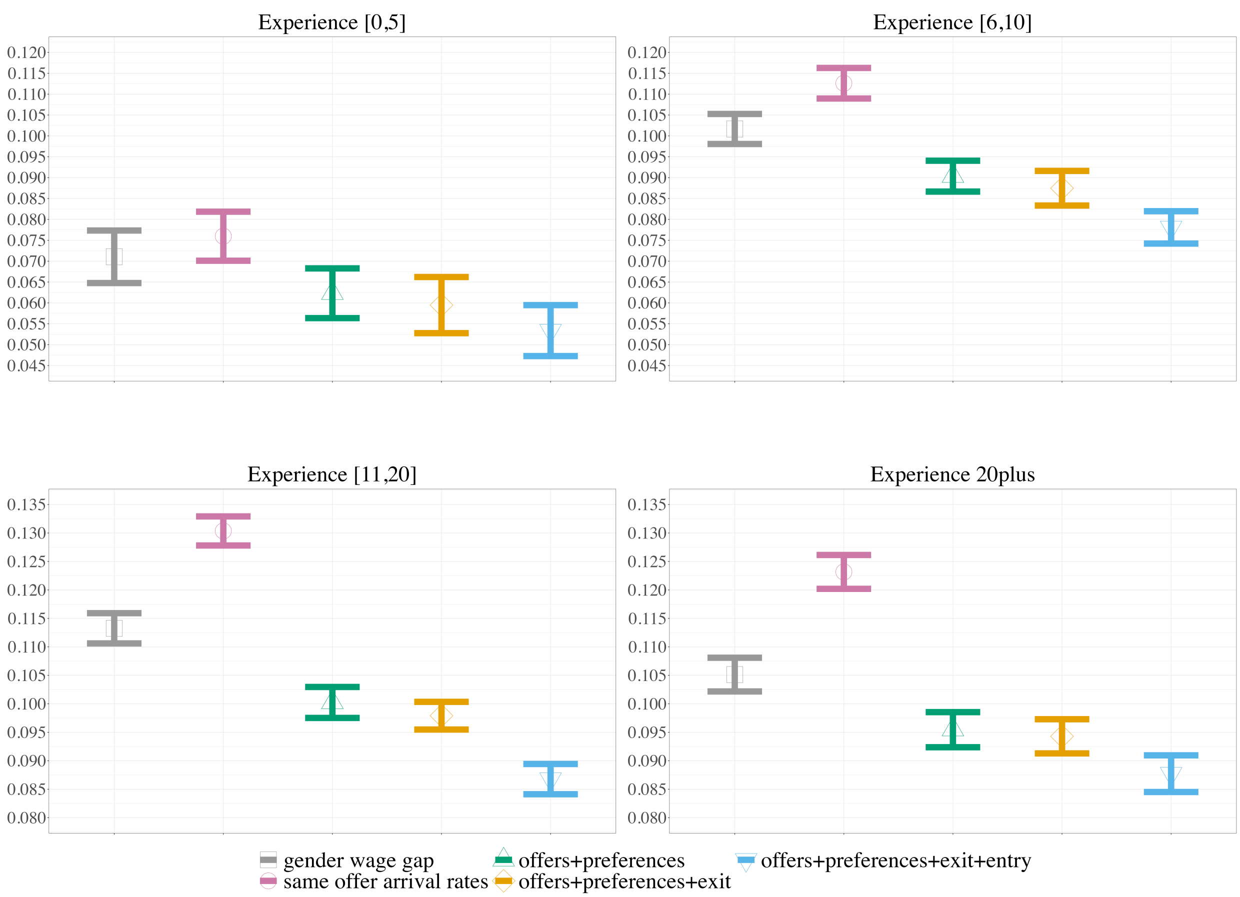
Notes: The figure compares the stationary worker-firm allocations between men and women. The allocations are obtained as described in Section 6. The x-axis represents the probability of working in firm class k for women of type l and characteristics x. The y-axis shows the same probability for men. The shapes on the graph indicate different levels of experience, with the darkness of colour intensifying to indicate higher-wage types and higher-paying firm classes. The dashed line represents the 45 degree line.

Figure 3: Worker Sorting and the Gender Wage Gap



Notes: The figure presents estimates of the gender wage gap for five counterfactual exercises, for the full sample. The counterfactuals simulate scenarios in which women have patterns in multiple mobility transitions as men do. Brackets indicate 95% confidence intervals obtained from 200 replications.

Figure 4: Worker Sorting and the Gender Wage Gap - By Experience

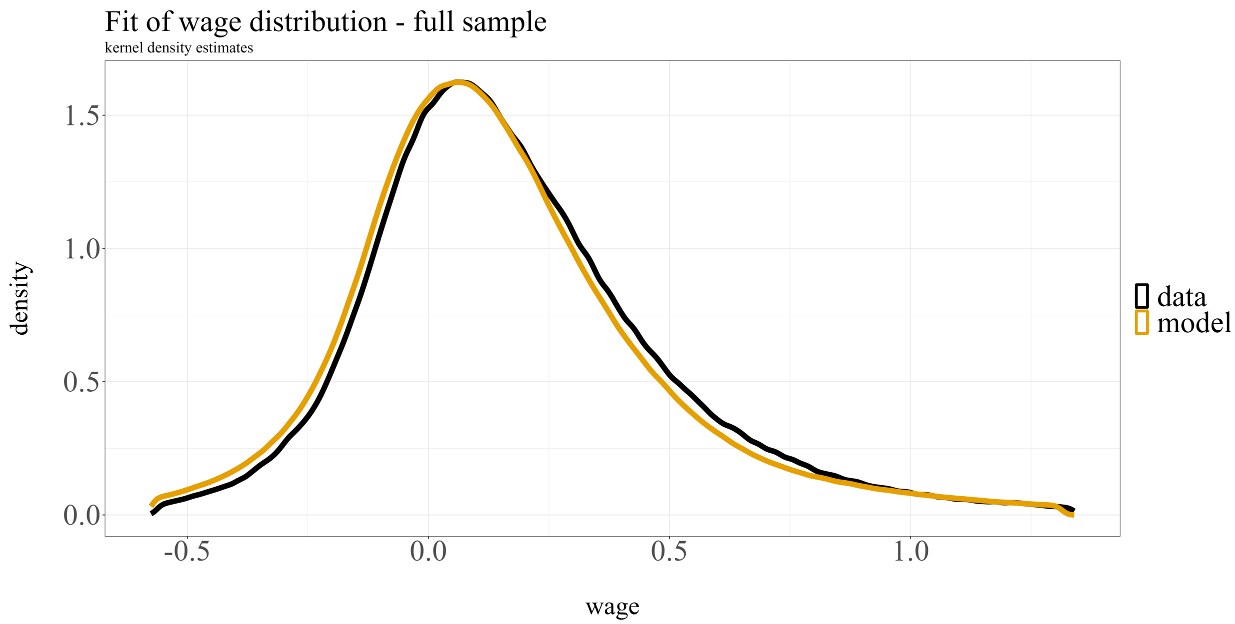


Notes: The figure presents estimates of the gender wage gap for five counterfactual exercises, by different combinations of experience groups. Each experience group includes both short and long tenured workers. The counterfactuals simulate scenarios in which women have patterns in multiple mobility transitions as men do. Brackets indicate 95% confidence intervals obtained from 200 replications.

Appendix

A Model Fit

Figure A1: Fit of wage distribution and gender wage gap



Notes: The figure shows the model fit. I simulate a dataset from the model estimates and compare the wage distributions of the simulated and actual data. In the top panel, the figure compares the full sample wage distributions. In the bottom panel, the figure compares the differences between the male and female wage distributions.

B Sample selection

Table B1: Firms' characteristics across selection samples

	Sample 1	Sample 2	Sample 3 (analysis)
N firms	100,424	48,667	25,925
N workers	73	145	250
N workers in Ile de France	27	53	92
% of women	46%	41%	43%
% of women in Ile de France	45%	40%	42%
% of women that are managers	18%	26%	30%
% of women in the board	36%	37%	39%
% of men	54%	59%	57%
% of men in Ile de France	55%	60%	58%
% of men that are managers	25%	31%	36%
% of men in the board	64%	63%	61%
Average Earnings	23,672	32,807	34,316
Median Earnings	22,363	30,545	31,688
Average hours	1,056	1,365	1,376
Median hours	1,106	1,457	1,498
Average hourly wages	20	23	24
Median hourly wages	19	21	22
% with part-time contracts	7%	10%	9%
% part-time and females	4%	6%	6%

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 in the DADS Panel.

Table B2: Workers' characteristics across selection steps

	Women							
	Earnings	Hours	Age	Experience	Tenure	% Part-Time	% Managers	N
Main job	20,678	1,225	41	17	5	29%	14%	1,378,721
Ile de France only	26,843	1,281	41	17	5	24%	24%	292,671
30+ days contract	27,030	1,289	41	17	5	24%	23%	289,443
+wages, +hours	27,021	1,292	41	17	5	23%	23%	287,211
Employed in Jan 2015	31,002	1,412	43	20	6	21%	26%	188,054
Never in Agriculture	31,005	1,412	43	20	6	21%	26%	187,926
Part-time & Full-time only	31,986	1,420	43	20	6	22%	27%	185,307
Aged 25-55	32,734	1,447	40	17	6	21%	29%	128,853
Never in seasonal/internship/domicile	34,359	1,493	41	17	6	19%	30%	117,314
Only in firms in DADS Postes	39,012	1,618	41	17	7	17%	34%	80,967
	Men							
	Earnings	Hours	Age	Experience	Tenure	% Part-Time	% Managers	N
Main job	27,820	1,343	40	17	5	14%	19%	1,441,594
Ile de France only	36,887	1,385	41	17	5	15%	31%	309,766
30+ days contract	37,177	1,396	41	18	5	14%	30%	306,025
+wages, +hours	37,152	1,400	41	17	5	14%	30%	303,896
Employed in Jan 2015	43,769	1,544	43	20	6	12%	34%	191,760
Never in Agriculture	43,793	1,544	43	20	6	12%	34%	191,482
Part-time & Full-time only	43,796	1,544	43	20	6	12%	34%	191,470
Aged 25-55	42,828	1,560	40	17	6	10%	35%	135,095
Never in seasonal/internship/domicile	45,794	1,623	41	18	6	8%	38%	122,378
Only in firms in DADS Postes	50,891	1,711	41	18	7	5%	42%	84,191

Notes: The table shows workers' characteristics at each selection step. The top panel compares female workers. The bottom panel compares male workers. The data source is DADS-Panel.

C Female shares

Table C1: Share of women by type, experience, and tenure

Tenure:	Short			Long		
Worker type:	Low-wage	Mid-wage	High-wage	Low-wage	Mid-wage	High-wage
Experience						
Experience 0-5	0.51	0.53	0.59	0.49	0.51	0.54
Experience 6-10	0.50	0.48	0.56	0.51	0.49	0.57
Experience 11-20	0.47	0.47	0.47	0.48	0.50	0.51
Experience 20+	0.45	0.43	0.43	0.47	0.47	0.47

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

D Estimates of the gender wage gap

Table D1: Worker Sorting and the Gender Wage Gap

	(1)	(2)	(3)	(4)	(5)
	Full Sample	First-period years of experience:			
		0-5	6-10	11-20	20+
Gender wage gap	0.110	0.071	0.102	0.113	0.105
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offer arrival rates	0.124	0.076	0.113	0.130	0.123
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences	0.098	0.062	0.090	0.100	0.095
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit	0.097	0.059	0.087	0.098	0.094
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit+entry	0.088	0.053	0.078	0.087	0.088
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)

Notes: The figure presents estimates of the gender wage gap for five counterfactual exercises. The counterfactuals simulate scenarios in which women have patterns in multiple mobility transitions as men do. Standard errors obtained from 200 replications are reported in parenthesis. The analysis assumes three worker types and five firm classes. The analysis assumes three worker types and fifteen firm classes.

Table D2: Alternative combination - 3 types and 5 classes

	(1)	(2)	(3)	(4)	(5)
	Full Sample	First-period years of experience:			
		0-5	6-10	11-20	20+
Gender wage gap	0.113	0.072	0.104	0.115	0.111
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offer arrival rates	0.125	0.078	0.117	0.130	0.120
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)
Same offers+preferences	0.104	0.063	0.094	0.105	0.103
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit	0.103	0.061	0.092	0.104	0.103
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit+entry	0.090	0.051	0.077	0.089	0.093
	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)

Notes: The figure presents estimates of the gender wage gap for five counterfactual exercises. The counterfactuals simulate scenarios in which women have patterns in multiple mobility transitions as men do. Standard errors obtained from 200 replications are reported in parenthesis. The analysis assumes three worker types and five firm classes.

Table D3: Alternative combination - 3 types and 10 classes

	(1)	(2)	(3)	(4)	(5)
	Full Sample	First-period years of experience:			
		0-5	6-10	11-20	20+
Gender wage gap	0.114	0.070	0.103	0.115	0.114
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offer arrival rates	0.134	0.072	0.113	0.138	0.143
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences	0.102	0.062	0.093	0.103	0.102
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit	0.102	0.059	0.090	0.102	0.102
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)
Same offers+preferences+exit+entry	0.090	0.052	0.079	0.089	0.091
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)

Notes: The figure presents estimates of the gender wage gap for five counterfactual exercises. The counterfactuals simulate scenarios in which women have patterns in multiple mobility transitions as men do. Standard errors obtained from 200 replications are reported in parenthesis. The analysis assumes three worker types and ten firm classes.

E K-means algorithm for firm clustering

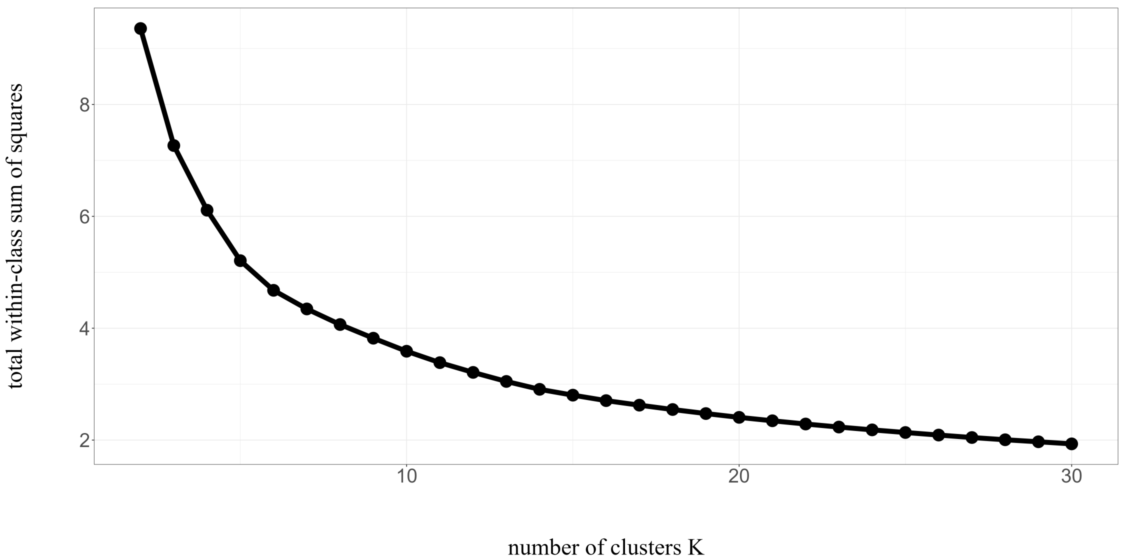
I partition the J firms into K distinct classes by solving a weighted k-means problem. As input characteristics, I use firms' male and female empirical cumulative distribution functions of log-hourly wages computed over five years and mean female shares, weighting by firms' average sizes. Log wages are 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 E1 shows the curve of the within-class sum of squares for a given number of classes K. I choose $K = 15$, seeking a balance between minimising total intra-class variation and ensuring sufficient observations.

To assess the stability of these clusters, I employ a bootstrap procedure (Jain and Moreau, 1987). For a fixed K, I generate new firm-level datasets through random sampling with replacement, maintaining the original dataset's size. Subsequently, I cluster the newly sampled data. Using the Jaccard similarity, I then identify the most similar cluster in the new clustering for each cluster in the original classification, repeating this process 100 times. Table E1 presents the average Jaccard similarity computed for each firm class 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 $K = 10$, in Table D2 and Table D3, respectively. Results are unaffected.

Figure E1: 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).

Table E1: Firm clusters stability

Firm class:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Avg Jaccard similarity	0.56	0.67	0.81	0.64	0.58	0.65	0.64	0.56	0.61	0.53	0.57	0.68	0.52	0.66	0.86

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.

F 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 $f(\theta^{(s+1)}) \geq f(\theta^{s})$.[17]

The log-likelihood relative to the likelihood in equation 1 is not linear in the mobility components of the parameter vector θ. In order to ease the maximization step, Lentz, Piyapromdee, and Robin (2023) consider a surrogate function that favours a straightforward maximization. In this Appendix, I describe the surrogate function of the log-likelihood that they use and go over the first-order conditions with respect to 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 is R.

17. 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.

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

F.1 Initial matching distribution

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

$$m_0\Big(l, k \,\Big|\, x\Big) = \frac{\sum_i p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\mathbb{1}\Big\{k_{j(i,1)} = k, x_{i1} = x\Big\}}{\sum_l \sum_k \sum_i p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\mathbb{1}\Big\{k_{j(i,1)} = k, x_{i1} = x\Big\}}$$

F.2 Wage Distribution Parameters

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

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

The wage segment of the expected log-likelihood writes:

$$W = \sum_i \sum_l p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \sum_k \sum_{t=1}^{T_i} \mathbb{1}\Big\{k_{j(i,t)} = k, x_{it} = x\Big\}\ln f(y_{it}|l, x, k)$$

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

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

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

F.3 Mobility Parameters

The mobility segment of the expected log-likelihood is:

$$Q = \sum_i \sum_l p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big)\Bigg\{\sum_t \sum_{k=0}^{K} \mathbb{1}\Big\{k_{j(i,t)} = k, x_{it} = x, s_{it} = 0\Big\}\ln M_{lxk\neg} + \\ \sum_t \sum_{k=0}^{K} \sum_{k'=0}^{K} \mathbb{1}\Big\{k_{j(i,t)} = k, k_{j(i,t+1)} = k', x_{it} = x, s_{it} = 1\Big\}\ln M_{lxkk'}\Bigg\}$$

Let k be the firm class of the current period and k' be the firm class of the subsequent period. Recall the parametric specification:

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

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

- $n^{(m)}_{lxk\neg} = \sum_i p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \sum_{t=1} \mathbb{1}\Big\{k_{j(i,t)} = k, x_{it} = x, s_{it} = 0\Big\}$
- $n^{(m)}_{lxkk'} = \sum_i p_i\Big(l \mid \theta^{(m)}, x_{i1}, C\Big) \sum_{t=1} \mathbb{1}\Big\{k_{j(i,t)} = k, k_{j(i,t+1)} = k', x_{it} = x, s_{it} = 1\Big\}$

Considering a given worker type l, the mobility segment of the expected log-likelihood becomes: $Q = \sum_{k=0}^{K} n_{lxk\neg}\Big(\theta^{(m)}\Big)\ln M_{lxk\neg} + \sum_{k=0}^{K}\sum_{k'=0}^{K} n_{lxkk'}\Big(\theta^{(m)}\Big)\ln M_{lxkk'}$

F.3.1 UE transition probabilities

For a given l, we obtain the M-step updating formulas for the unemployment to employment transition probabilities by deriving with respect to $\psi_{lxk'}$ the following segment of the expected log-likelihood: $n^{(m)}_{lx0\neg} \ln M_{lx0\neg} + \sum_{k'=1}^{K} n^{(m)}_{lx0k'} \ln M_{lx0k'}$. Plugging in the parametric specification:

$$n^{(m)}_{lx0\neg} \ln\Big(1 - \sum_{k'=1}^{K} \psi_{lxk'}\Big) + \sum_{k'=1}^{K} n^{(m)}_{lx0k'} \ln(\psi_{lxk'})$$

Taking derivatives, we obtain:

$$\psi_{lxk'}^{(m+1)} = \frac{n_{lx0k'}^{(m)}}{n_{lx0\neg}^{(m)} + \sum_{k'=1}^{K} n_{lx0k'}^{(m)}} \tag{F.3}$$

F.3.2 EU and JTJ transition probabilities

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

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

Under the parametric specification above, this expected log-likelihood segment is not linear in the job-to-job transition parameters. I, therefore, consider the minorising function proposed by Lentz, Piyapromdee, and Robin (2023). For $k \in \{1, \ldots, K\}$, we can write:

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

In other words, a l-type worker stays in the same firm class k if she does not receive an offer/layoff or if she receives an offer from k' but prefers to stay in k. In order to build the minorising function, we first notice that the following equality holds.

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

Exploiting the concavity of the logarithm, the following inequality holds.

$$\ln(M_{lxk\neg}) = \ln\left(1-\delta_{lxk}-\sum_{k'=1}^{K}\lambda_{lxk'}+\sum_{k'=1}^{K}\lambda_{lxk'}(1-P_{lxkk'})\right) \geq$$
$$\frac{1-\delta^{(s)}_{lxk}-\sum_{k'=1}^{K}\lambda^{(s)}_{lxk'}}{M^{(s)}_{lxk\neg}}\ln\left(\frac{1-\delta_{lxk}-\sum_{k'=1}^{K}\lambda_{lxk'}}{1-\delta^{(s)}_{lxk}-\sum_{k'=1}^{K}\lambda^{(s)}_{lxk'}}\;M^{(s)}_{lxk\neg}\right)+$$
$$\sum_{k'=1}^{K}\frac{\lambda^{(s)}_{lxk'}(1-P^{(s)}_{lxkk'})}{M^{(s)}_{lxk\neg}}\ln\left(\frac{\lambda_{lxk'}(1-P_{lxkk'})}{\lambda^{(s)}_{lxk'}(1-P^{(s)}_{lxkk'})}M^{(s)}_{lxk\neg}\right) \equiv \ln(\underline{\mathrm{M}}_{lxk\neg})$$

The inequality becomes an equality if $\lambda^{(s)}_{lxk'} = \lambda_{lxk'}$, $\delta^{(s)}_{lxk} = \delta_{lxk}$, and $P^{(s)}_{lxkk'} = P_{lxkk'}$. That is, $\ln(\underline{\mathrm{M}}_{lxk\neg})$ minorizes $\ln(M_{lxk\neg})$. We can thus consider $\ln(\underline{\mathrm{M}}_{lxk\neg})$ instead of $\ln(M_{lxk\neg})$ and the MM algorithm maximizes:

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

Given $\theta^{(m)}$ obtained at the m-step of the EM algorithm, I update δ, γ, λ by maximising $H(M|\theta^{(m)})$ using an iterative procedure described below. First, define:

- $\tilde{n}^{(s)}_{lxkk'} = n^{(m)}_{lxk\neg}\frac{\lambda^{(s)}_{lxk'}(1-P^{(s)}_{lxkk'})}{M^{(s)}_{lxk\neg}}$ the predicted number of l-type stayers that receive an offer from k' but prefer to stay in k.
- $\hat{n}^{(s)}_{lxk} = n^{(m)}_{lxk\neg}\frac{1-\delta^{(s)}_{lxk}-\sum_{k'=1}^{K}\lambda^{(s)}_{lk'}}{M^{(s)}_{lxk\neg}}$ the predicted number of l-type stayers that stay because they receive no offer and no layoff.

We plug $\tilde{n}^{(s)}_{lxkk'}$ and $\hat{n}^{(s)}_{lxk}$ into $H(M|\theta^{(m)})$ and update γ_{lxk} by maximising the segment:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'}\ln\frac{\gamma_{lxk}}{\gamma_{lxk}+\gamma_{lxk'}} + \sum_{k=1}^{K}\sum_{k'=1}^{K} n^{(m)}_{lxkk'}\ln\frac{\gamma_{lxk'}}{\gamma_{lxk}+\gamma_{lxk'}}$$

With a simple change of indices:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'}\ln\frac{\gamma_{lxk}}{\gamma_{lxk}+\gamma_{lxk'}} + \sum_{k=1}^{K}\sum_{k'=1}^{K} n^{(m)}_{lxk'k}\ln\frac{\gamma_{lxk}}{\gamma_{lxk}+\gamma_{lxk'}}$$

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxk'k}\right)\,\ln\gamma_{lxk}\;-\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxk'k}\right)\ln(\gamma_{lxk}+\gamma_{lxk'})$$

Following Hunter (2004) we note that:

$$-\ln(\gamma_{lxk}+\gamma_{lxk'})\geq 1-\ln(\gamma^{(s)}_{lxk}+\gamma^{(s)}_{lxk'})-\frac{\gamma_{lxk}+\gamma_{lxk'}}{\gamma^{(s)}_{lxk}+\gamma^{(s)}_{lxk'}}$$

Therefore, we can consider the following segment:

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxk'k}\right)\,\ln\gamma_{lxk}\;+\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxk'k}\right)\left[1-\ln(\gamma^{(s)}_{lxk}+\gamma^{(s)}_{lxk'})-\frac{\gamma_{lxk}+\gamma_{lxk'}}{\gamma^{(s)}_{lxk}+\gamma^{(s)}_{lxk'}}\right]$$

With an additional change of indices and with simple algebra, we update γ_{lxk}:

$$\gamma^{(s+1)}_{lxk}=\frac{\sum\limits_{k'=1}^{K}(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxk'k})}{\sum\limits_{k'=1}^{K}\left(\frac{\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxkk'}+\tilde{n}^{(s)}_{lxk'k}+n^{(m)}_{lxk'k}}{\gamma^{(s)}_{lxk}+\gamma^{(s)}_{lxk'}}\right)}\tag{F.4}$$

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

$$\begin{aligned}&\sum_{k=1}^{K}\hat{n}^{(s)}_{lxk}\ln\Big(1-\delta_{lxk}-\sum_{k'=1}^{K}\lambda_{lxk'}\Big)+\sum_{k=1}^{K}\sum_{k'=1}^{K}\tilde{n}^{(s)}_{lxkk'}\ln(\lambda_{lxk'})\\&+\sum_{k=1}^{K}n^{(m)}_{lxk0}\ln(\delta_{lxk})+\sum_{k=1}^{K}\sum_{k'=1}^{K}n^{(m)}_{lxkk'}\ln(\lambda_{lxk'})\end{aligned}$$

Taking derivatives with respect to δ_{lxk}:

$$n^{(m)}_{lxk0}\Big[1-\sum_{k'=1}^{K}\lambda_{lxk'}\Big]=\Big(n^{(m)}_{lxk0}+\hat{n}^{(s)}_{lxk}\Big)\delta_{lxk}\tag{F.5}$$

Taking derivatives with respect λ_{lxk}:

$$\frac{1}{\lambda_{lxk'}}\sum_{k=1}^{K}\left(\tilde{n}^{(s)}_{lxkk'}+n^{(m)}_{lxkk'}\right)=\sum_{k=1}^{K}\left(\frac{\hat{n}^{(s)}_{lxk}}{1-\delta_{lxk}-\sum\limits_{k'=1}^{K}\lambda_{lxk'}}\right)\tag{F.6}$$

Explicit δ_{lxk} in F.5, plug it into F.6, and update $\lambda_{lxk'}$ and δ_{lxk} as follows:

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

$$\delta_{lxk}^{(s+1)} = \frac{n_{lxk0}^{(m)}\left(1 - \sum_{k'=1}^{K} \lambda_{lxk'}^{(s+1)}\right)}{n_{lxk0}^{(m)} + \hat{n}_{lxk}^{(s)}} \tag{F.8}$$

For given value of $\theta^{(m)}$, the sequence $H(M|\theta^{(m)})$ increases at each iteration step s of the MM algorithm. It is thus not strictly necessary to wait for convergence. The algorithm can be stopped at any time. I iterate the MM algorithm 200 times before it delivers the updated values $\delta^{(m+1)}$, $\gamma^{(m+1)}$, and $\lambda^{(m+1)}$.