Labour market concentration and gender gaps

Johanne Bacheron* Sylvie Blasco† Eva Moreno - Galbis‡ Jeremy Tanguy§

Abstract

This paper analyses how labour market concentration affects gender inequalities in wages, hirings, and working conditions. While theoretical models predict that firms will be able to extract a monopsony rent from workers who have lower geographical mobility, very specific skills, or specific working conditions' requirements, there is limited empirical evidence on this topic. Using French matched employer-employee data together with data on working conditions and a new definition of commuting zones that incorporates gender differences in mobility, we find that concentration in a given commuting zone and occupation increases the gender wage gap and decreases the share of women among new hires, but has limited effect on the gender gap in working conditions. Women with children and women of childbearing age are particularly affected by the increase in firms' monopsonistic power.

Keywords: labour market concentration, monopsony, gender gap, wages, working conditions
JEL codes: J31, J42, L13

*Aix Marseille Univ, CNRS, AMSE.

†Université Caen Normandie, CREM, IZA.

‡Aix Marseille Univ, CNRS, AMSE. The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference: ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University - A*MIDEX.

§Université Savoie Mont Blanc, IREGE.

1 Introduction

In a strict sense monopsony refers to a market structure with a single buyer confronted in a market by many sellers (see Robinson (1969)). However, monopsony power can be actually exercised by any employer facing an upward sloping labour supply curve. When the number of competitors of a firm on a market is very large, we converge towards a perfect competitive situation where the firm acts as a "wage taker". In contrast, as the number of competitors in its operating market falls, the firm increases its monopsonistic power, and therefore its bargaining position when negotiating wages with potential employees. The larger the bargaining power of the firm, the larger the ability of the firm to extract rents from the employment relationship.

One limitation in part of the literature on local labour markets is the definition of the workforce as an homogeneous entity. The literature generally defines local labour markets as commuting zones or the intersection between the commuting zone and the industry/occupation. However, this definition implicitly assumes that all workers residing in that local area have access to all the jobs proposed in that market, and have homogeneous preferences concerning both wage and non wage characteristics of the job, such as geographical mobility [1].

It has been shown in the literature that this is not the case, in particular women have a lower willingness to commute than men, and they are willing to accept lower wages in exchange for a reduced commuting time (Le Barbanchon et al., 2021). Since their utility loss of geographical mobility is larger, they implicitly have a smaller local labour market than the size objectively defined by the commuting zone or the intersection between the commuting zone and the industry/occupation, firms have thus a stronger monoposonistic power over them since they have smaller outside employment opportunities. We take this into account by defining gender-specific commuting zones using the algorithm provided by INSEE and census data on women's and men's commuting patterns, which allows us to have a more accurate measure of concentration for each gender. However, with a same level of concentration, women may still face lower wages due to firms being more selective in presence of asymmetric information on the labour market. Employers may prefer workers with more work experience (Bassanini et al., 2021), or for a same level of experience, they may discriminate against women because they see them as less committed to work due to household-related responsibilities. This can translate to hiring discrimination (Becker et al., 2019), and lower wages (Xiao, 2021). It may disproportionately impact women with children or young women. Therefore we also distinguish by parental status to investigate the mechanisms behind our findings.

We use a simplified version of the circular model proposed in Gautier and Zenou (2010), with

[1]Bhaskar and To (1999) develop a model of monopsonistic competition in which workers have heterogenous preferences over some non-wage characteristics of potential jobs. In their particular model this preference is over a measure of distance to the job (closer is better). In their model, workers facing equal wage offers accept the closer offer.

a search-matching in the labour market and wage bargaining. The labour market is composed by men and women workers. Women are assumed to have a larger utility loss from the distance between their preferred geographical location and the actual geographical location of the job. The model predicts that the range of acceptable geographical locations for a job is narrower for women than for men. As a result, the wage that women manage to bargain is also lower. This effect is reinforced if we additionally assume that, as a result of the lower geographical mobility, women have lower outside employment opportunities. Finally, it is interesting to underline that, the narrower the range of acceptable geographical locations, the lower the outside options and therefore the narrower should be the range of working conditions over which the worker can bargain.

The model's assumptions and predictions are tested using French data. In a first stage, we test the main assumption of the model. Due to prevailing social norms concerning housework sharing within heterosexual couples, women are traditionally less geographically mobile than men. This hypothesis is made by the model. We exploit the longitudinal dimension of the Panel Tout Salariés - *Échantillon Démographique Permanent (EDP)* (*i.e.* Panel All Salaried workers and Census) 2009-2019 to compare geographical mobility of men and women, and find that women commute significantly less than men, in particular when they have children.

In a second stage we test the predictions of the model concerning wages. We do so by exploiting the French DADS salariés database from 2009 to 2019. We then define local labour markets as the intersection between the gender-specific commuting zones and occupations, and compute the Herfindahl-Hirschman Index (HHI) by local labour market. Because of their smaller commuting zones, concentration is higher for women. We then estimate if higher labour market concentration is associated with a larger gender wage gap, controlling for establishment-level productivity and product market concentration through the inclusion of establishment-by-time fixed effects. To circumvent endogeneity issues we instrument the HHI in each occupation-commuting zone using the employment-weighted average HHI within the same occupation across other commuting zones. We find that concentration negatively affects women's wages 30% more that men's using IV, and twice as much using OLS. When firms are constrained from adjusting wages, in presence of high minimum wages for example, their increase in bargaining power can translate into a change in hiring strategies. We show that an increase in labour market concentration decreases the share of women among new employees, an in particular that of young women.

In a third stage, we study if this lower geographical mobility of women translates not only into relative worse pecuniary conditions but also into worse non-pecuniary working conditions. Using the Working Conditions Survey (WCS) 2013, 2016, 2019 we define a range of indicators of non-pecuniary working conditions. We replicate the regression taking as dependent variables our indicators of working conditions. Our results are sensitive to the specification chosen. We observe an increase in the gender gap in job intensity, development opportunities, scheduling, and in the

non-pecuniary index in some specifications, but no significant effect in others.
Our paper contributes to an emerging literature analysing the consequences of labour market concentration on labour market outcomes (wages, employment, hirings).[2] There are two streams in the monopsony literature, each one adopting an alternative approach but both being actually complementary. The first stream of monopsony literature focuses on the elasticity of labour supply to the individual firm. If workers have a high labour supply elasticity, then firms pay them more to get them to stay. If workers have low labour supply elasticity, then firms can exploit their monopsonistic power and extract rents by paying workers a low wage (below their marginal productivity). This stream of literature generally finds low elasticities of labour supply and interprets this as evidence for firm-level monopsony power.
The first study using data at the establishment level in order to estimate the labour supply elasticity appears to be Sullivan (1989). He estimates the supply elasticity of nurses directed toward individual hospitals to be in the range 1.3–3.8. Staiger et al. (2010) argue that the caseload was endogenous in the period studied by Sullivan (1989). Staiger et al. (2010) exploit then a legislative increase in the wage for nurses at U.S. Veterans Administration hospitals to estimate the elasticity of labour supply of these nurses to be 0.1.[3]
Falch (210) estimates the elasticity of teacher labour supply to individual schools in Norway to be about 1.4 and is in the range 1.0-1.9. Ransom and Oaxaca (2010) focus on the elasticity of labour supply of teachers in public school districts in Missouri. They leverage the exogenous variation in prenegotiated district salary schedules to instrument for actual salary. They estimate a labour supply elasticity of about 3.7, pointing to the significant market power of school districts. Moreover, elasticity of supply is smaller for women suggesting a gender gap of 8%.
Using US employer-employee data, Webber (2016) estimates that the labour supply elasticities are 1.09 and 0.94 for men and women, which leads to 3.3 percent lower earnings for women. Using survival analysis and a large linked employer-employee data set for Germany, Hirsch et al. (2010) estimate that labour supply elasticities are small (1.9–3.7). Women's labour supply is less elastic than men's and this implies that at least one third of the gender pay gap may be explained by

[2]The literature has analysed other impacts of labour market concentration in the economy. Based in US data Autor et al. (2020) and Barkai (2020) find that a decline in labour shares driven by the increased concentration of the labour market with the appearance of superstar firms. Exploiting firm-level data for the US economy since 1955, De Loecker et al. (2020) document the evolution of market power, markups and profitability. They estimate that average markups start to rise in 1980 from 21% above marginal cost to 61% now. The average profit rate increased from 1% to 8%. Moreover, they find a reallocation of market share from low to high markup firms.

[3]Staiger et al. (2010)'s findings are in great contrast with Matsudaira (2014) who exploit the approval of a state minimum staffing law for nurses in California to estimate firm level elasticity of labour supply for nurse aides in the long-term care (nursing home) industry. He finds that facilities initially out of compliance with the new law did not have to raise their wage offers relative to their competitors in order to hire more nurses. This is consistent with the perfect competition hypothesis rather than with monopsony.

profit maximizing monopsonistic employers that exploit women's lower labour supply elasticity.[4] Finally, in a more recent work, Dube et al. (2020) use a double machine learning estimator applied to a large dataset of scraped MTurk tasks to isolate plausibly exogenous variation in rewards and estimate in this way the elasticity of labour supply. They also reanalyse data from five MTurk experiments that randomized payments to obtain corresponding experimental estimates. Both approaches yield uniformly low labour supply elasticities, around 0.1.

Estimates on the supply elasticity provided by papers in this stream of literature are all very small and suggest considerable monopsony power in a variety of settings. They also imply a considerable size of monopsony rents (difference between the marginal product and the actual wage paid by employers).

Our paper relates to the second stream of literature, which proposes a complementary approach with a different mechanism at play. Papers in this stream of literature measure market concentration in local, sectoral or occupational labour markets. The implicit idea of this approach is that workers do not engage in geographically wide-ranging job searches. Therefore, firms will be able to pay lower wages when there are few competing firms in the market (*i.e.* high concentration in the labour market), whereas they will have to pay higher wages the larger the number of firms present in the market (*i.e.* the lower the degree of concentration of the labour market) since workers will bring employers into competition in order to negotiate higher wages. In this stream of literature it is the buyer-side market power caused by concentration that pushes wages down because in the presence of few firms in the labour market (*i.e.* high labour market concentration) workers will not be able to drive firms into competition when negotiating their wage.

Papers exploiting the buyer-side market power approach are more recent than those estimating the elasticity of the upward-sloping firm labour supply curve as measure of monopsonistic power. Additionally, while most of the former literature estimating labour supplies elasticities considers particular labour markets (*i.e.* teachers, nurses, retail workers, etc.), this stream of literature considers the whole labour market, which allows to have a better view on how widespread labour market power is and the impact on wages. Azar et al. (2022) exploit data from the employment website CareerBuilder.com to calculate labour market concentration for over 8,000 geographic-occupational labour markets in the US. They estimate that moving from the 25th percentile to the 75th percentile in concentration is associated with a 5% (OLS) to 17% (IV) decline in posted wages, suggesting that concentration increases labour market power and puts downward pressure in wages. Rinz (2022) combines comprehensive US administrative data on firms and individuals with demographic information obtained from surveys to consider distributional effects of local industrial

[4] Applying duration models that account for unobserved worker heterogeneity, Hirsch and Janhn (2015) exploit a large administrative employer-employee German dataset and find that firm-labour supply elasticity is significantly smaller for immigrants with respect to natives (1.136 vs. 1.360). This differential predicts 7.7 log points wage penalty for immigrants.

concentration on earnings and inequality within and between demographic groups. He focuses in the period 1976-2015 and exploits time variation across markets in the magnitude of changes in local industrial concentration. He finds that following the reduction in average concentration, the 90/10 earnings ratio was 6% lower and earnings 1% higher in 2015 than they would have been if local concentration had remained as its 1976 level. All demographic groups experience increases in inequality when concentration increases. Using manufacturing plant-level U.S. Census data during 1978-2016, Benmelech et al. (2022) provide evidence that wages are lower in local labour markets in which employers are more concentrated. This negative relationship between employer concentration and wages increases over time and it is particularly strong when unionisation rates are low. They also find that the link between productivity growth and wage growth is stronger when markets are less concentrated.

The closer papers to ours are Bassanini et al. (2021), Marinescu et al. (2021) and Arquie and Bertin (2023). They both exploit linked employer-employee French administrative data to analyse the impact of labour market concentration on labour market outcomes. Bassanini et al. (2021) consider the period 2010-2017 and estimate the elasticity of stayer's wages with respect to labour market concentration. Their estimations range between -0.0185 and -0.0230. Marinescu et al. (2021) focus on 2011-2015. They define concentration for new hires. More precisely, the authors compute the Herfindahl-Hirschman Index for new hires at the occupation-commuting zone-quarter level. They find that a 10% increase in labour market concentration decreases hires by 3.2% and their hourly wage by 0.5% which is consistent with predictions of a monopsony model. Arquie and Bertin (2023) focus on the period 2000-2019 and analyse the consequences of concentration on wage distribution. They do not only consider the impact of employer's concentration on overall inequality (between jobs) but also on within-firm and between-firm inequality. They estimate that concentration increases wage inequality by undercutting relatively more the bargaining power of the lowest earners.

The contribution of our paper to this second stream of literature is threefold. First of all, we propose a standard circular model à la Salop (1979) with matching frictions to illustrate how differences in preferences towards geographical mobility reduce wages through the reduction in the range of acceptable locations for a job and in the outside employment opportunities. Second, since women are less geographically mobile than men we compute gendered commuting zones to study the impact of labour market concentration on the gender wage gap. Third, in contrast with most existing literature, we study the relationship between monopsony power and two additional outcomes : the gender gaps in hiring and the (non-pecuniary) working conditions. To our knowledge only Bassanini et al. (2023), Qiu and Sojourner (2023) and Meiselbach et al. (2022) focus on the impact of labour market concentration on non-wage attributes. Bassanini et al. (2023) find that higher concentration negatively affects job security whereas the two other papers find a negative effect of concentration on employer-provided health in the U.S.

The paper is organised as follows. Section 2 presents the theoretical framework. Databases, variables and descriptive statistics are presented in Section 3. The econometric strategy is described in section 4. Section 5 analyses the impact of market concentration on the gender gap in wages, working conditions and hirings. Section 6 studies the economic mechanisms behind the results. Section 7 concludes.

2 Theoretical Framework

2.1 Labour market flows

We consider a simplified version of the circular model proposed in Gautier and Zenou (2010) and retaken in Moreno Galbis (2020).[5] There is a continuum of risk neutral workers and firms. In both cases the mass is normalized to 1. Time is continuous and workers live forever. A proportion p of the population is composed by males, M, and $(1-p)$ by females, F. Both are assumed to be identical apart from the fact that females are less geographically mobile. We represent this difference across genders by assuming that women have a larger utility loss associated with geographical mobility.

Following Salop (1979) and in line with the model proposed in Gautier and Zenou (2010), workers' and firms' heterogeneity over geographical location is modelled by means of a circle. Workers and firms are assumed to be uniformly distributed over this circumference C of length 1. This is the space of possible locations. We denote by $0 < x_{ij} < 1/2$ the distance between a worker located in i and a firm located in j. It is assumed that workers do not change their residence over their lifetime.

Workers may be employed or unemployed. All unemployed workers search for a job and we assume that there is no on-the-job search. Let $u_k(i)$ be the number of type-k unemployed workers, $k = M, F$ (or equivalently the unemployment rate of type k-workers) located in i. At each moment in time, a firm can either have a filled position or an open vacancy. We denote $v(j)$ the number of vacancies (or equivalently the vacancy rate) located at j.

As shown in Gautier and Zenou (2010), the uniform distribution of workers over the circle implies that $u_k(i) = u_k \ \forall i \epsilon C$. In this case, there exists a stationary equilibrium with a uniform and unique distribution of vacancies at all locations $v(j) = v \ \forall j \epsilon C$.

As in the standard search-matching models, individuals choose reservation wages by comparing the values of employment and unemployment. But here workers must also decide on the width of the range of acceptable geographical locations and firms must decide their location given the geographical preferences of workers.

[5] Staiger et al. (2010) is an application of Salop (1979)'s model of competition around a circle without matching frictions in the labour market.

Search is random and the number of contacts between workers and firms is given by $M = M(u_M + u_F, v)$, which is increasing in its arguments, concave, exhibits constant returns to scale and satisfies Inada conditions. The probability of finding a job equals $m(\theta) = M/(u_M + u_F)$, where $\theta = v/(u_M + u_F)$ stands for the labour-market tightness. Evidently the probability of finding a job is increasing in θ. In contrast, the probability of filling a vacancy, $M/v = q(\theta) = m(\theta)/\theta$, is decreasing in θ. The contact rate of a vacancy with a type-k worker is then given by $q(\theta) \cdot u_k/(u_M + u_F)$.
For a worker, a match will occur if and only if the location of the firm enters within her acceptable range, that is:

$$Match_{u_k \to v} = m(\theta) \cdot 2\hat{x}_k$$

A match is the product of a random contact rate $m(\theta)$ and an acceptation rule $2\hat{x}_k$. All workers randomly get job contacts at the rate $m(\theta)$. Once the contact takes place, workers must decide to apply for this job or not, depending on the geographical location associated with the job with respect to the worker's location, which would correspond to the most preferred location of the worker. The term $\hat{x}_k$ is multiplied by 2 because each worker considers the distance from both sides of the circle with respect to her position.
Applying an analogous reasoning we define the rate at which vacancies match with a type-k worker as:

$$Match_{v \to u_k} = \frac{u_k}{u_M + u_F} \frac{m(\theta)}{\theta} \cdot 2\hat{x}_k$$

After advertising their jobs, firms are contacted by workers. They will only offer a job to type-k workers located within a distance $x < \hat{x}_k$.
Denoting e_k the employment level (or rate) of workers of type k and δ the exogenous job destruction rate, the steady state equilibrium flows for k-type workers can be written as: $\delta e_k = 2\hat{x}_k m(\theta) u_k$
Since the labour force is normalized to unity, $p = u_M + e_M$ and $(1-p) = u_F + e_F$, we can write the equilibrium flow equations for each nativity group as follows:

$$\begin{aligned} \delta(p - u_M) &= 2\hat{x}_M m(\theta) u_M \Rightarrow u_M = \frac{\delta p}{\delta + 2\hat{x}_M m(\theta)} \\ \delta(1 - p - u_F) &= 2\hat{x}_F m(\theta) u_F \Rightarrow u_F = \frac{\delta(1-p)}{\delta + 2\hat{x}_F m(\theta)} \end{aligned}$$

2.2 Workers' behavior

As in Gautier and Zenou (2010), workers change jobs over their lifetime but not their residence so that the difference between desired (or most preferred) and accepted geographical location changes stochastically over time. As a result, the average difference between most preferred and actual geographical location of the job is the same for all workers of type k over their lifetime.

Outside employment opportunities are represented by R (it includes unemployment benefits, social aids, wealth, alternative employment offers, etc). For simplicity we are going to assume that R is identical across genders. The only source of difference between males and females will be the disutility, $\tau_k(x)$, associated with geographical distance between the most preferred location of the worker and the actual location of the job. We will assume that the utility loss is larger for women because they dislike relatively more geographical mobility.

Defining U_k as the steady-state expected discounted lifetime utility of an unemployed worker of type k and $E_k(x, w_k)$ as the steady state expected discounted lifetime utility of an employed worker of type k in a job implying a difference with respect to preferred geographical location equal to x and earning a wage w_k, we can define:

$$rU_k = R + m(\theta)\Big[2\int_0^{\hat{x}_k} [E_k(x, w_k) - U_k]\Big]dx \tag{1}$$

where $r\epsilon(0, 1)$ is the discount rate, and $\hat{x}_k$ is the maximum distance between the most preferred and actual work location a worker is willing to accept (beyond $\hat{x}_k$ all jobs will be turned down by the unemployed workers). When a worker of type k is unemployed today, her instantaneous utility equals R. She then meets vacancies at the rate $m(\theta)$ but only a fraction $\hat{x}_k$ of the vacancies are acceptable for the worker. When a worker accepts a job paying w_k and at a distance x from her most preferred work location, she obtains a wealth increase of $[E_k(x, w_k) - U_k]$. U_k does not depend on x because search is random so that firms cannot sort workers depending on their most preferred work location.

The asset value for an employed worker who is employed in a job located at distance x from the worker's most preferred location equals:

$$rE_k(x, w_k) = w_k - \tau_k(x) - \delta(E_k(x, w_k) - U_k) \tag{2}$$

where $\tau(x)$ stands for the utility loss associated with distance between the worker's preferred location and the actual geographical location of the job. We refer to disutility rather than to transport cost since disutility can vary across individuals depending on their preference for geographical mobility, while the transport cost is in principle similar across genders. We assume the disutility to be linearly increasing in x, that is $\tau_k'(x) > 0$. The employed worker obtains an instantaneous utility $w_k - \tau_k(x)$ from the job but can lose her job with probability δ and experience a reduction in wealth equal to $(E_k(x, w_k) - U_k)$.

2.3 Firms' behavior

Let y be the productivity of a worker and γ denote the firm's search cost per unit of time. Since we assume constant returns to scale production, profits do not depend on firm size so we can consider all vacancies to be single worker firms. The expected discounted lifetime utility of a firm with a filled job is given by:

$$rJ_k(w_k) \;\; = \;\; y - w_k - \delta(J_k - V) \tag{3}$$

where $y - w_k$ corresponds to the instantaneous profit and $\delta(J_k - V)$ to the loss if the job is destroyed. Since productivity does not depend on the distance between the most preferred and the actual job location all employed workers are, in terms of productivity, identical from the point of view of the firms.

The expected discounted lifetime utility of a firm with a vacancy is given by:

$$V \;\; = \;\; -\gamma + 2\frac{m(\theta)}{\theta}\Big(\frac{u_F}{u_F + u_M}\int_0^{\hat{x}_F}(J_F - V)dx + \frac{u_M}{u_F + u_M}\int_0^{\hat{x}_M}(J_M - V)dx\Big) \tag{4}$$

Every period the firm pays an advertisement cost γ. The contact with a worker takes place at rate $\frac{m(\theta)}{\theta}$ and the firm can meet a female with probability $\frac{u_F}{u_F+u_M}$ or a male with probability $\frac{u_M}{u_F+u_M}$. The female will accept the job if the proposed work location fall within the distance $\hat{x}_F$ and the male accepts the job if the distance is below $\hat{x}_M$.

2.4 The steady state equilibrium

A (steady-state) labour market equilibrium is a tuple that consists of wages, a maximum acceptable distance between most preferred and actual geographical location, unemployment levels and labour-market tightness $(w_F^*, w_M^*, \hat{x}_F^*, \hat{x}_M^*, u_F^*, u_M^*, \theta^*)$. Given the matching technology, in this equilibrium all agents (workers and firms) maximize their respective objective function. labour market tightness is determined by a free-entry condition, wages by Nash bargaining, and maximum geographical distance by an indifference condition between the value of unemployment and the value of employment at this maximum acceptable distance. Finally, unemployment and vacancy levels follow from equilibrium labour-market tightness and a steady-state condition on unemployment.

Labour demand

Firms open vacancies until no more profit can be obtained. At the equilibrium all rents are exhausted and the free entry condition, $V = 0$, applies. Equation (3) becomes then:

$$J_k = \frac{y - w_k}{r + \delta} \tag{5}$$

Combining this equation with equation (4) in the presence of the free entry condition leads to:

$$\frac{\gamma\theta(r+\delta)}{2m(\theta)} = \frac{u_F(y - w_F)\hat{x}_F + u_M(y - w_M)\hat{x}_M}{u_F + u_M} \tag{6}$$

Therefore $\frac{\partial \theta}{\partial \hat{x}_k} > 0$. If workers become less picky concerning acceptable geographical location, firms create more jobs since they have more chances to fill them.

Wage determination

Firms do not observe workers' most preferred geographical location or more precisely their x. As in Gautier and Zenou (2010), workers and firms only bargain over observable factors. The total surplus of the match is then given by: $\Omega_k = \left[E_k^d - U_k^d + J_k - V\right]$. Following the Nash bargaining process, the surplus is shared in constant proportions according to the respective bargaining powers of workers and firms. Denoting by $0 < \eta < 0$ the bargaining power of the worker, the equilibrium wage is given by:[6]

$$w_k = \eta y \frac{(r + \delta + 2m(\theta)\hat{x}_k)}{r + \delta + \eta 2m(\theta)\hat{x}_k} + (1 - \eta) R \frac{(r + \delta)}{r + \delta + \eta 2m(\theta)\hat{x}_k} \tag{7}$$

where $\frac{\partial w_k}{\partial R} > 0$ and $\frac{\partial w_k}{\partial \hat{x}_k} > 0$.[7] The higher the unearned income, the higher the wage. Similarly, the larger the range of acceptable geographical locations $\hat{x}_k$, the better the outside options and therefore the higher the bargained wages. Since θ, δ, y, R and r are the same for males and females, if $\hat{x}_M > \hat{x}_F$ males should be able to bargain higher wages.

Range of acceptable geographical locations

We must finally determine the maximum distance a worker is willing to accept between her most preferred geographical location and the geographical location actually proposed by the job, *i.e.* $\hat{x}_k$. Formally, $\hat{x}_k$ is implicitly defined by the geographical distance that makes the worker indifferent between being employed or remaining unemployed, *i.e.* $E_k(\hat{x}_k, w_k) = U_k$, which leads to:[8]

$$\hat{x}_k = -\frac{r + \delta}{2m(\theta)} + \frac{1}{2m(\theta)} \sqrt{(r + \delta)^2 + \frac{4m(\theta)(r + \delta)(w_k - R)}{\mu_k}} \tag{8}$$

where $\frac{\partial \hat{x}_k}{\partial \mu_k} < 0$, $\frac{\partial \hat{x}_k}{\partial R} < 0$ and $\frac{\partial \hat{x}_k}{\partial w_k} > 0$.[9] The lower the disutility, μ_k, generated by the distance between the most preferred job location and actual location, the larger the range of geographical locations the worker is willing to accept. Similarly, the lower the level of unearned income, R, the larger the range of acceptable work conditions. In contrast, higher wages are associated with a larger range of acceptable work conditions. This is in line with conclusions drawn from the wage equation (17).

[6] See Appendix A for more details in the derivation process.

[7] $\frac{\partial w_k}{\partial R} = (1-\eta)\frac{r+\delta}{r+\delta+\eta 2m(\theta)\hat{x}_k} > 0$ and $\frac{\partial w_k}{\partial \hat{x}_k} = \frac{2m(\theta)\eta(1-\eta)(r+\delta)(y-R)}{(r+\delta+\eta 2m(\theta)\hat{x}_k)^2} > 0$.

[8] See Appendix A for more details

[9] $\frac{\partial \hat{x}_k}{\partial \mu_k} = -\frac{(r+\delta)(w_k - R)}{\mu_k^2 \sqrt{(r+\delta)^2 + \frac{4m(\theta)(r+\delta)(w_k-R)}{\mu_k}}} < 0$ as far as $w_k > R$ which is always the case if the worker accepts the job.

Similarly $\frac{\partial \hat{x}_k}{\partial R} = -\frac{(r+\delta)m(\theta)}{\mu_k} \cdot \frac{1}{m(\theta)\sqrt{(r+\delta)^2 + \frac{4m(\theta)(r+\delta)(w_k-R)}{\mu_k}}} < 0$ and $\frac{\partial \hat{x}_k}{\partial w_k} = \frac{(r+\delta)m(\theta)}{\mu_k} \cdot \frac{1}{m(\theta)\sqrt{(r+\delta)^2 + \frac{4m(\theta)(r+\delta)(w_k-R)}{\mu_k}}} > 0$.

As shown by Gautier and Zenou (2010) a single crossing condition for equations (17) and (23) applies since at $\hat{x}_k = 0$ the concave function (17) has a higher value of w_k than the convex function (23), implying the unique pair solution $(\hat{x}_k, w_k)$ exists.

2.5 Model's prediction

According to our model, $\frac{\partial \hat{x}_k}{\partial R} < 0$, $\frac{\partial \hat{x}_k}{\partial \mu_k} < 0$ and $\frac{\partial \hat{x}_k}{\partial w_k} > 0$. We can also easily verify that $\frac{\partial \hat{x}_k^2}{\partial R \partial \mu_k} > 0$ confirming that the reduction in the range of acceptable locations following a reduction in R will be more important for individuals bearing a high disutility associated with the distance between most preferred and actual job location.

Prediction 1: If women have higher utility loss associated with geographical mobility, $\mu_F > \mu_M$, the range of acceptable geographical locations will be narrower for them.

Prediction 2: This lower range of acceptable geographical locations will weaken their bargaining position decreasing their wages.

For identical outside employment opportunities, R, the range of acceptable geographical locations will be narrower for females because of their higher utility loss in case of geographical mobility. If there is a global decrease in outside employment opportunities, R, the range of acceptable geographical locations will be narrower for both males and females.
If instead of assuming that workers and firms negotiate only over wages, we assume that they negotiate also over other working conditions, we easily deduce that the narrower the range of acceptable geographical locations, the lower the outside options and therefore the lower the bargaining position of the worker to negotiate good working conditions, including wages. Therefore, reduced geographical mobility decreases both bargained wages and working conditions, since workers with reduced geographical mobility cannot bring into competition as many employers as workers with high geographical mobility. While there does not exist a pure monopsony situation, reduced geographical mobility improves the relative bargaining position of firms over workers.

3 Data

3.1 Databases

We combine a wide variety of databases. While they are not freely accessible, any researcher can request access to them through the Secure Data Access Centre (CASD). First, we use French Administrative employee-employers data named Déclaration annuelle de données sociales (DADS-Postes) collected by the INSEE (Institut Nationale de la Statistique et des Etudes Economique) between 1995 and 2019. It provides us with individual level data on workplace location, wages, hours worked, occupation, industry, gender, and age. For multi-establishment firms, the data provides establishment identifier and location, which allows us to distinguish between workers employed in different establishments of the same firm. This administrative dataset covers all French private and public sector workers, however the worker identifier changes every year. It is thus an exhaustive repeated cross-section of workers that allows us to identify individual workers and their primary source of income (*i.e.* the job providing them with the most income during a given year). We use this dataset to measure the level of labour concentration by commuting zone and occupation.

We also use the panel dimension of these employers-employees data, the DADS Panel - EDP (*i.e.* Échantillon Démographique Permanent, that corresponds to the population Census), which is a representative sample of the exhaustive DADS based on the date of birth and which covers 4% of the DADS. This dataset is not exhaustive, making it unreliable to construct concentration indices. However, its longitudinal dimension will be exploited to study the relationship between labour market concentration and wages and will allow us to include individual fixed effects. The combination of the DADS Panel with Census data allows us to control for the age and the number of children of individuals.

The FH (Fichier Historique) provides information on the characteristics of job-seekers, and, in particular, their reservation wage, their maximum commuting distance or the type of jobs they are looking for. At the beginning of each unemployment spell, when they register to the unemployment agency (Pôle Emploi), job-seekers must indicate the minimum gross wage they are willing to work for, and the maximum distance they are willing to commute to work each day (one way). This will allow us to illustrate how the willingness to commute differs between unemployed men and women. The reservation wage can be reported on an hourly, daily, or annual basis, and we convert it to a monthly reservation wage. The maximum commuting accepted can be expressed in minutes or kilometers. When provided in minutes, we convert it to kilometers under the assumption of an average speed of 35 kilometers per hour.

Lastly, we use the Working Conditions Survey 2013, 2016, 2019 to study the link between concentration and the gender gap in working conditions. This survey aims to obtain a concrete description of work, its organisation and its conditions from various angles: schedules, work rhythms, physical efforts or risks incurred, hardship, work organisation, safety, cooperation, conflicts, etc. The survey has been conducted for the past 40 years: every seven years until 2005 and every three years since 2013, and allows us to analyse the evolution of work conditions. Since 2013 there is a panel dimension and it is possible to match every worker to their employer when exploiting the data through the Secure Data Access Centre.

3.2 Sample selection

We focus on the period 2009-2019, during which occupations are consistently defined. We only keep private sector employees and, in line with Marinescu et al. (2021) we exclude state-sponsored workers, apprentices, interns, workers in non-governmental organisations, in the art industry, museums, sport clubs, agriculture, unions and at home. We consider individuals above 23 years old, so as to avoid student jobs, and below 62 years old, which corresponds to the average retirement age in France. Because France has a binding minimum wage (and compliance is very high), we discard for every year the 2% lowest wages and, to avoid upward outliers, we also discard the 2% highest wages.[10] Because partial time jobs are more frequent among women, we cannot focus exclusively on full time workers, since there will be selection issues with the sample of women. To avoid this problem, we focus on hourly wages.

In line with Arquie and Bertin (2023), if a firm owns several establishments in the same labour market, we consider that the jobs of all these establishments belong to one and unique entity, which we consider as a unique employer. Employees of all establishments owned by the same firm within the same labour market are considered as being employed by the same entity. We keep only firms with at least two employees in a given local labour market as firms with only one employee might be very specific ones.

We define local labour markets as the intersection of commuting zones and occupations. As remarked in Bassanini et al. (2021) employees change jobs across industry borders and workers in different occupations within a given industry do not compete for the same jobs. Therefore, local labour markets should not be considered at the intersection between a commuting zone and a sector. Also, when defining the labour market by commuting zone-sector, we cannot introduce plant-by-time fixed effects, since a given plant operates in one single industry and geographical area, so this FE would be colinear to any measure of labour market concentration defined with respect to an industry at the geographical area. The point is that introducing plant-by-time fixed effects allows us to control for productivity changes.

3.3 Variables

3.3.1 Definition of the labour market Herfindahl-Hirschman Index (HHI)

Labour market concentration is measured through the employment Herfindahl-Hirschman Index (HHI). The HHI is defined as the sum of the shares of employment in a given market. We have defined local labour market at the intersection of occupation at the four digit level and commuting zone. Due to the division of household responsibilities, women typically experience

[10] Arquie and Bertin (2023) remove observations whose log annualized real earnings are more than 5 standard deviations away from a predicted wage computed using a linear model including socio-demographic controls. This procedure leads them to exclude between 3% and 5% of all observations each year.

shorter commuting times than men and are often willing to accept lower wages for jobs located closer to their places of residence. We thus compute a measure of concentration specific to the male and the female labour markets by using gender-specific commuting zones. In order to construct them, we use the algorithm used by INSEE to define the 2020 labour market areas [11], based on the analysis of commuting between the different areas, and apply it separately on women's and men's commuting patterns provided in the 2019 census data.[12] The methodology used to create these gender-specific local labour market areas is presented in Appendix B.

A firm's f labour market share in occupation j and commuting zone c in year t will be equal to:

$$s^e_{j,c,f,t} = \frac{N_{j,c,f,t}}{\sum_f N_{j,c,f,t}} \tag{9}$$

where $N_{j,c,f,t}$ represents the number of workers employed by firm f in occupation j, commuting zone c in year t. The employment-HHI in the corresponding local labour market defined by occupation j and commuting zone c in year t is then:

$$HHI^e_{j,c,t} = \sum_f (s^e_{j,c,f,t})^2 \times 100 \tag{10}$$

By definition the HHI is always between 0 and 100. When it is equal to 100, that means that a single employer employs all workers in the labour market. An HHI between 15 and 25 is indicative of a moderately concentrated market and above 25 of a highly concentrated market (see guidelines of the American Department of Justice and Federal Trade Commission). Since women have a smaller range of acceptable locations for their jobs and therefore smaller commuting zones, labour market concentration in their commuting zone is likely to be higher

An HHI based on employment is reasonable approximation of the index of labour market concentration that is relevant for wage determination in a stationary search and matching model with granular search as ours, where concentration affects wages by changing workers' outside options (see Jarosch et al., 2021). In a non-stationary environment, downsizing of firms may have a positive share in the stock of employment in a local labour market, whereas their hirings are zero, so that they do not contribute to creating outside options for workers in that labour market. This argument is used by Bassanini et al. (2021), Bassanini et al. (2023) or Marinescu et al. (2021) to justify using new hires to compute the HHI index. In our case, our model predicts that outside employment opportunities are driven by geographical mobility of workers, rather than by firms

[11] The algorithm LabourMarketAreas (R package) has been developed by Eurostat. It was initially presented in Coombes and Bond (2008) and is an evolution of the classical methodology of the "Travel-To-Work Areas" (TTWA), defined in Coombes et al. (1986). This algorithm led in France to the definition of 306 different labour market areas in 2020, against 321 previously with another method.

[12] The algorithm exploits individual information on place of residence and place of work available in the 2019 census data – professional mobility detail file (*Recensement de la population 2019, fichier détail - Mobilités professionnelles*), provided by INSEE.

dynamics. In any case, since firm dynamics is likely to affect almost symmetrically both men and women our results should not be strongly modified by focusing on employment or new hires when computing the HHI.[13]

3.3.2 Working conditions

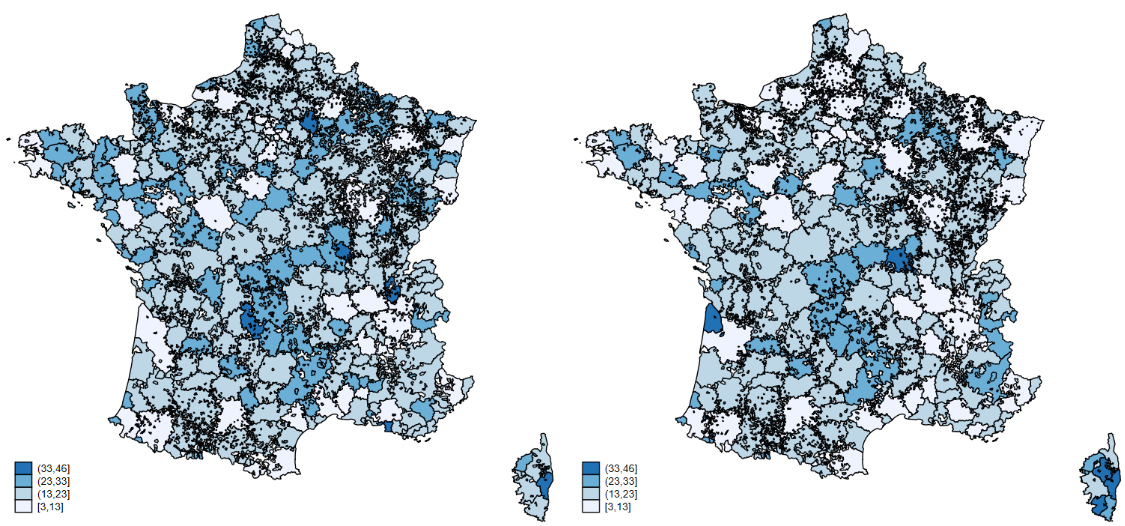
We define the eleven non-pecuniary working conditions indicators as follows: (i) *Learning new things* equals unity if the individual declares learning new things on the job; (ii) *Autonomy* results from the addition of four discrete variables defined between 0 and 1 and increasing with the degree of autonomy of the worker in solving problems, choosing methods, deciding the speed of working and determining the quality of the good or service; (iii) *Support* results from the addition of two dummy variables which equal unity if the individual declares receiving support from colleagues and/or from their manager; (iv) *Stability* takes a maximum value of three if the individual has a unlimited duration contract, believes that he has weak probability of loosing his job in the next six months, and has high seniority; (v) *Development opportunities* results from the addition of a dummy variable for receiving training from the employer and a dummy for having prospects for career advancement; (vi) *Physical safety* is an indicator which increases if the exposure to physical risks for the individual at work is lower. It results from the addition of seven discrete indicators defined between 0, when the individual is exposed to the corresponding physical risk, and 1 when the individual is not exposed to the risk. We consider the following physical risks: vibrations, loud noises, smoke, exposure to chemical products, working in a painful position, moving loads and/or implementing repetitive actions; (vii) *Psychological safety* equals 0 if the individual has to deal with angry customers and 1 if the individual does not have to deal with them; ($viii$) *Scheduling* is equal to four for an individual who does not have to work at night, on Saturdays, on Sundays, or during their free time; (ix) *Commuting* is between 0 and 1 such as the higher the score, the lower the commuting time, (x) *Flexibility* is equal to five if the individual is not being controlled for his working time, does not have his work arrangement changed regularly, has fixed timetables, has no shifts, is able to take easily 1-2 hours off during the day, and is able to take a break when he wants; (xi) *Intensity* is equal to three for an individual who does not have to work at high speed, does not have deadlines and has enough time to do his work. To summarise, the higher the value of each of these indicators the better the working conditions. Lastly, we construct a Non-Pecuniary Index that is equal to the average of these eleven indicators.

[13] In the presence of downsizing/upsizing firms in a local market both men and women will see their outside employment opportunities decrease/increase.

3.4 Descriptive statistics

Figure 1 presents the gender-specific commuting zones, as constructed with the algorithm provided by INSEE for labour market areas, and the corresponding labour concentration indices. We can see that women's commuting zones are, consistently with our predictions, smaller than those of men since they are less mobile than them. In total in Metropolitan France, we estimate that there are 271 different commuting zones for women and 216 for men. Consequently, women face higher levels of labour market concentration. Higher levels of concentration can be observed also in low populated areas, particularly along the "empty diagonal" stretching from the North-East to the South-West of the country. A potential endogeneity issue may arise from the fact that low density areas can be both those with high concentration and lower wages (and can also be those where gender gaps are higher). However, we address this concern in the next section by incorporating a set of fixed effects that should capture these variations.

Figure 1: Labour market concentration by gender-specific commuting zone



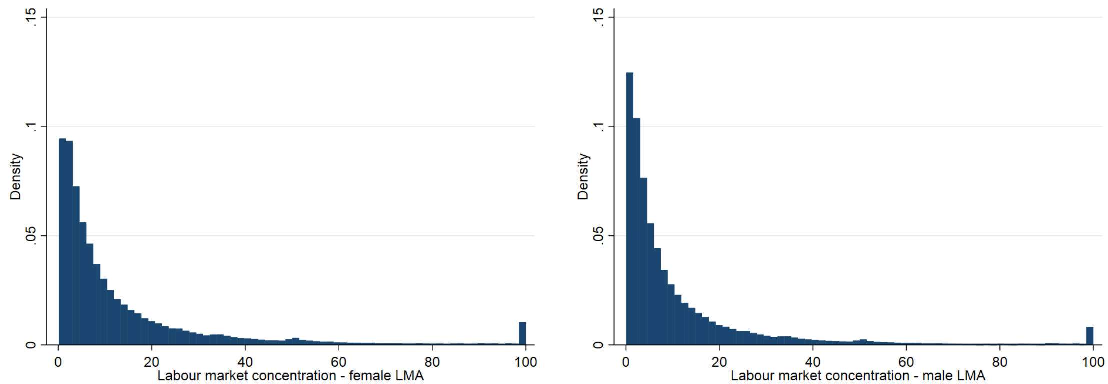
(a) Female-specific commuting zones

(b) Male-specific commuting zones

Source: Panel Tout Salariés-EDP

The difference in labour market concentration between genders can also be seen in Figure 2, which presents the distribution of the HHI for female and male local labour markets. We can see that men are more likely to face levels of concentration very close to zero. The average HHI for women in the sample is equal to 14, whereas for men it is equal to 12, while the median of the index is equal to 6.8 and 5.4 respectively.

Figure 2: Distribution of labour market concentration by gender



(a) HHI for women

(b) HHI for men

Source: Panel Tout Salariés - EDP

Tables 1 and 2 show descriptive statistics on labour market outcomes and working conditions respectively. We can see that labour market are not very concentrated in France, with an average HHI of 13 when taking the individual-level data, while the median HHI is equal to 6, meaning that half of the individuals in the sample face a concentration index that is below 6. This table also shows that unemployed women that are present in the FH-DADS have a monthly reservation wage which is almost 300€ lower than their male counterparts, and that they are willing to commute on average 7km less[14]. Using the Panel Tout Salariés, we estimate that employed women have an average commuting distance which is 40% lower than men[15].

We present also in Table C.1 in Appendix the results of a linear regression of the average commuting time of employed individuals by gender and parental status, which is also an important determinant of mobility. We can see that even after controlling for the commuting zone and year, women commute on average 11km less than men. On average, parents commute less than non-parents, and this is true for both genders, but the difference is slightly higher for men. However, adding individuals fixed effects shows that men commute on average 0.5km less after the first child, whereas women commute 2km less. Regarding job-seekers (Table C.2 in Appendix), we observe the same phenomenon, with a maximum commuting distance accepted which is lower for women on average, and which decreases with the first child, in particular for women.

These results are consistent with the assumption of our model according to which women are less geographically mobile than men. Based on these descriptive evidence, we will also use the parental status to investigate the mechanisms behind our findings in Section 6.

[14]This trade-off between commuting and wages in the French context has been studied by Le Barbanchon et al. (2021), who find that the differences in the willingness to commute account for 14% of the residual wage gap.

[15]We use the distance between the centroid of the municipality of residence and the centroid of the municipality of work to compute the commuting distance for employed individuals.

Table 1: Summary statistics - labour outcomes

	Min	Max	Median	Mean	SD	N
Panel Tout Salariés - EDP						
HHI	.10655	100	6.079	13.160	18.795	65061400
Male	0	1	1	.560	.496	6529444
French nationality	0	1	1	.880	.324	6529444
Age	24	61	40	40.811	10.300	6529444
Experience	0	49	11	11.628	6.345	6529444
Hourly wage	6.594	40.051	11.701	13.585	5.798	6529444
Number of children	0	12	0	.628	.899	6529444
No diploma	0	1	0	.131	.337	5757641
Lower secondary education	0	1	0	.376	.484	5757641
Upper secondary education	0	1	0	.214	.410	5757641
Short-cycle tertiary education	0	1	0	.161	.367	5757641
University diploma	0	1	0	.118	.323	5757641
Commuting distance - women	0	417.744	7.866	15.933	36.552	2742301
Commuting distance - men	0	541.690	10.478	26.404	61.954	3502971
FH - DADS						
Reservation wage - women	0	4983.876	1398.397	1558.916	471.6484	475858
Reservation wage - men	0	7495	1500	1820.824	782.0136	415195
Maximum commuting distance - women	0	200	20	23.10502	14.97654	447299
Maximum commuting distance - men	0	200	30	30.63558	21.3926	378253

Regarding working conditions, Table 2 shows the score of employed men and women in each of the 11 indicators and in the non-pecuniary index, the higher the score and the better the working condition. We can see that on average men's working conditions are slightly better than that of women. They score higher in all but four points: they have less autonomy in their jobs, they tend to work more on weekends or holidays, they have higher commuting time, and are more subject to physical risks than women.

Table 2: Summary statistics - non pecuniary working conditions

	Women				Men			
	Min	Max	Mean	N	Min	Max	Mean	N
Learning new things	0	1	0.773	29046	0	1	0.788	22241
Autonomy	0	4	2.026	29083	0	4	1.938	22269
Support	0	2	1.410	29083	0	2	1.483	22269
Stability	0	2.741	1.741	29083	0	2.833	1.759	22269
Development opportunities	0	2	0.994	29052	0	2	1.101	22246
Physical safety	0	7	5.210	29081	0	7	4.611	22264
Scheduling	0	4	2.868	29083	0	4	2.813	22269
Commuting	0	1	0.906	28629	0	1	0.894	22069
Flexibility	0	5	3.440	29081	0	5	3.614	22266
Intensity	0	3	1.966	29083	0	3	2.022	22269
Psychological safety	0	1	0.542	24026	0	1	0.592	14868
Non-Pecuniary Index	.572	3.205	2.112	23769	.544	3.189	2.162	14715

Source: Working Conditions Survey
Survey weights are used.

Tables C.1, C.2 and C.3 in Appendix show the correlation between the hourly wages or the different working conditions and the labour market concentration, separately by gender and without any controls. Each point on the graph corresponds to a commuting zone. We can see a clear decreasing relationship between the log hourly wages and the HHI for both men and women. Regarding working conditions, we can see a flat (albeit slightly decreasing for women) relationship between the labour market concentration and the non-pecuniary index. The HHI seems to affect negatively the likelihood to feel supported in the workplace for both men and women, decreases the job stability, the psychological safety, and make scheduling more difficult only for women. The possibility to learn new things increases with the HHI for women, whereas the psychological safety increases for men. It is also interesting to note that individuals tend to commute less when the concentration increases. The other working conditions do not seem correlated with the labour market concentration.

4 Econometric Strategy

For reasons related to the sharing of tasks within couples, women tend to seek work closer to home than men and are willing to accept a wage discount to reduce their commuting time (Jacob et al., 2019; Le Barbanchon et al., 2021). This suggests that the relevant local labour market for women is smaller than for men. Instead of assuming identical commuting zones for men and women, a common practice in the literature, we recognize that differences in the impact of labour market concentration on genders may actually reflect variations in the size of their respective relevant labour markets. We thus use the gender-specific labour market areas that we computed based on women's and men's commuting patterns.

We estimate:

$$\begin{aligned}\log(w_{i,j,o,c,t}) =& \alpha \log(HHI_{o,c,t}) + \beta \log(HHI_{o,c,t}) * Female_i + \mathbf{X}'_{i,j,o,c,t}\gamma \\ &+ \mu_i + \mu_{oc} + \mu_{jt} + \mu_t^M + \mu_t^F + \varepsilon_{i,j,o,c,t}\end{aligned} \tag{11}$$

where $w_{i,j,o,c,t}$ is the hourly wage of individual i, in establishment j, in occupation o, in commuting zone c, and in year t. The commuting zone here differs between women and men and sometimes overlaps. As we cannot include occupation-gender-specific commuting zone fixed effects, we try the specification by either including the occupation-female commuting zones or the occupation-male commuting zones (μ_{oc}) and make sure that the results do not depend on the geographical area chosen. $\mathbf{X}$ is a vector of individual time varying controls. μ_{jt} are establishment-by-time fixed effects. In the more demanding specification, we include individual fixed effects (μ_i) and gender-year fixed effects (μ_t^M and μ_t^F).

A threat to identification is the existence of time-varying market-specific variables that are correlated with concentration and affect wages. For example, a decline in market dynamism is likely to lead to a reduction in the number of jobs and to an outward migration of young workers towards more dynamic labour markets. Moreover, according to our theoretical framework, bargained wages are influenced by productivity, labour market tightness, outside employment opportunities and the share of acceptable geographical locations. To take into account this time-varying market-specific variables we control for establishment-level productivity and product market concentration by including establishment-by-time-fixed effects.
In spite of our efforts to control for observable and unobservable confounders through the introduction of control variables and fixed effects, endogeneity issues remain a concern. More precisely, a biased productivity shock benefiting relatively larger firms could affect both concentration and the progression in the gender wage gap. Typically, if the gender gap is larger in large firms, a biased productivity shock pushing small firms out of the market and pushing up the size of large firms will drive an increase in concentration and in the gender wage gap. To circumvent endogeneity issues we propose a instrumental variable strategy similar as the one used in Azar et al. (2022), Rinz (2022) and Arquie and Bertin (2023). We instrument for the HHI in each local labour market (*i.e.* occupation-commuting zone), using the employment-weighted average HHI within the same occupation across other commuting zones, excluding the one considered. This instrument provides variation in market concentration driven by national-level changes in the occupation, rather than local changes in that particular local market. This approach helps mitigate endogeneity concerns in cases of asymmetric productivity shocks across commuting zones.

The instrument for labour market concentration in sector j, commuting zone c, in period t equals:

$$HHI_{-c,j,t} = \frac{\sum_v (N_{j,v,t} \cdot HHI_{j,v,t})}{\sum_v N_{j,v,t}} \tag{12}$$

where v represents all commuting zones but c, j indexes occupation and t years. Again, employment in occupation j, in all commuting zones except from c v, in period t is denoted by $N_{j,v,t}$.
Alternatively, we also instrument the HHI with the average of $log(1/F)$ in other commuting zones for the same occupation and time period, where F refers to the number of firms in the market. $log(1/F)$ is less likely to be endogenous than the instrument based on HHI from other labour markets, as it does not depend on market shares. In line with the instrument based on the HHI, $log(1/F)$ provides us with variation in market concentration that is driven by national-level changes in the occupation, and not by changes in that particular local market. In particular, $log(1/F)$ should be independent from productivity shocks in the local labour market, which is the main confounding factor in the baseline OLS regression.

5 Results

5.1 Hourly wages

We present in the section the estimations on the effect of labour market concentration on the gender gap in hourly wages, first with the OLS method (Table 3), and then the instrumental variable approach with the two different instruments (Tables 4 and 5). All columns include control variables at the individual level and establishment-by-year fixed effects. From column (2), we include commuting zone-occupation fixed effects. Because male and female commuting zones can overlap, we cannot include gender-specific commuting zone fixed effects. We thus control in columns (2) to (4) by female commuting zone - occupation fixed effects, and in columns (5) to (7) by male commuting zone - occupation fixed effects, and check that the choice of the geographical area does not affect our results. We then further control for individual, and gender-year fixed effects.
We can see first in Table 3 that controlling for commuting-zone $\times$ occupation explains a significant part of the gender wage gap: the coefficient associated with the dummy female decreases by almost 40% once we account for commuting-zone/occupation. Consistently with the literature, we find a negative and significant relationship between labour market concentration and hourly wages, at least once we account for commuting zone/occupation, even though the coefficients are small. The most demanding specification suggests that a 10% increase in the HHI index decreases the average hourly wage by 0.013%. Lastly, we can see that our main coefficient of interest, the interaction

between being a woman and the index of labour market concentration, that is positive without individual fixed effects, becomes negative once we include them, which suggests there might be unobserved heterogeneity at the individual-level that explains both the concentration and the gender gap. The linear regression suggest that the negative effect of labour market concentration is almost twice as high for women as it is for men.

To deal with potential endogeneity issues, we replicate the estimations using instrumental variables, first by instrumenting the HHI by the employment-weighted average HHI across other commuting zones (in Table 4), and by the average of $log(1/F)$ in other commuting zones (in Table 5), with F the number of firms in the market. The results are consistent with what we found with OLS, although there is a change in the magnitude of the coefficients. The IV estimates suggest a higher effect of concentration on the average wages, as a 10% increase in the HHI decreases the average wage by about 0.05-0.07% in our preferred specification with establishment-year, commuting zone-occupation and individual fixed effects. The effect on the gender wage gap is lower with the IV approach, as we find that concentration decreases women's wage by about 30-35% more than men's. These results are consistent with our hypothesis that the wage penalty is higher for women with the same level of labour market concentration.

Table 3: Effect of HHI on hourly wages - OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.127***	-0.0786***			-0.0781***		
	(0.000579)	(0.000577)			(0.000570)		
Log HHI	0.0107***	-0.00370***	-0.00130***	-0.00132***	-0.00469***	-0.00161***	-0.00163***
	(0.000198)	(0.000379)	(0.000363)	(0.000363)	(0.00366)	(0.000360)	(0.000360)
Female × Log HHI	0.00765***	0.00230***	-0.000985**	-0.000998**	0.00199***	-0.000965**	-0.000989**
	(0.000239)	(0.000232)	(0.000333)	(0.000334)	(0.000230)	(0.000327)	(0.000328)
N	3693707	3681612	4079812	4079812	3685711	4084071	4084071
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes	No	No	No
Male CZ × occupation FE	No	No	No	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	No	Yes	Yes
Gender-year FE	No	No	No	Yes	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience. Since men and women have different commuting zones which overlap, we first control for female commuting zones-occupation fixed effects in columns (2) to (4), and then for male commuting zones-occupation fixed effects in columns (5) to (7).

Table 4: Effect of HHI on hourly wages - IV HHI instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.139***	-0.0777***			-0.0781***		
	(0.000720)	(0.000844)			(0.000869)		
Log HHI	0.00902***	-0.00913***	-0.00577***	-0.00604***	-0.00793***	-0.00570***	-0.00595***
	(0.000260)	(0.00132)	(0.000996)	(0.000997)	(0.00134)	(0.000994)	(0.000995)
Female × Log HHI	0.00143***	0.00246***	-0.00202***	-0.001988**	0.00239***	-0.00175***	-0.00173***
	(0.000323)	(0.000332)	(0.000474)	(0.000476)	(0.000334)	(0.000470)	(0.000472)
N	3693705	3681611	4079811	4079811	3685710	4084070	4084070
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes	No	No	No
Male CZ × occupation FE	No	No	No	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	No	Yes	Yes
Gender-year FE	No	No	No	Yes	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience. Since men and women have different commuting zones which overlap, we first control for female commuting zones-occupation fixed effects in columns (2) to (4), and then for male commuting zones-occupation fixed effects in columns (5) to (7).

Table 5: Effect of HHI on hourly wages - IV Log 1/F

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.0851***	-0.0655***			-0.0658***		
	(0.000836)	(0.001000)			(0.00101)		
Log HHI	0.0726***	0.00653***	-0.00674***	-0.00720***	0.00693***	-0.00698***	-0.00743***
	(0.000300)	(0.00144)	(0.00122)	(0.00122)	(0.00144)	(0.00121)	(0.00121)
Female × Log HHI	-0.0127***	-0.00474***	-0.00268***	-0.00276**	-0.00485***	-0.00239***	-0.00246**
	(0.000387)	(0.000418)	(0.000547)	(0.000530)	(0.000416)	(0.000539)	(0.000523)
N	3690848	3678905	4076814	4076814	3682901	4080947	4080947
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes	No	No	No
Male CZ × occupation FE	No	No	No	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	No	Yes	Yes
Gender-year FE	No	No	No	Yes	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience. Since men and women have different commuting zones which overlap, we first control for female commuting zones-occupation fixed effects in columns (2) to (4), and then for male commuting zones-occupation fixed effects in columns (5) to (7).

We replicate these estimations by using a measure of HHI based on new hires to check if our results hold when we change the definition of concentration and present them in Tables D.1 to

D.3 in Appendix. We can see that the coefficient associated with the concentration index looses its significance, but we still observe an increase in the gender wage gap with higher levels of concentration.

5.2 Working conditions

If we assume that workers and firms negotiate not only over wages, but also over other working conditions, we easily deduce that the narrower the range of acceptable geographical locations, the lower the outside options and therefore the lower the bargaining position of the worker to negotiate good working conditions, including wages. Therefore, reduced geographical mobility decreases both bargained wages and working conditions, since workers with reduced geographical mobility cannot bring into competition as many employers as workers with high geographical mobility. Alternatively, we can argue that, if workers consider reduced geographical mobility as an important working condition, they may be willing to sacrifice other working conditions, pecuniary and/or non pecuniary, in order to keep this reduced geographical mobility. Therefore, reduced wages may simply denote that workers sacrifice pecuniary working conditions in exchange of reduced geographical mobility. These worse pecuniary working conditions may be associated with worse non-pecuniary conditions, depending on how much workers are willing to sacrifice in order to have a low commuting time.

To test this hypothesis we estimate Equation 11 using as a dependent variable our working condition indicators. Because the number of observations in the Working Condition Survey is relatively low, we want to ensure that at least 10 individuals are present in every commuting zone-occupation, we thus restrict the sample to these cases. Tables 6 and 8 present the results of the OLS and IV estimations for each indicator and for the index of non-pecuniary working conditions. The results are sensitive to the specification chosen. Using weighted least squares, we obtain that the non-pecuniary index decreases with concentration, with no differential effect by gender, whereas job intensity increases for women compared to men (Table 6). The first instrumental variable regression shows a positive relationship between concentration and learning new things on the job (Table 7), whereas the second one suggests a negative relationship on average between the HHI and job stability, a positive relationship with physical safety, and a higher gender gap in terms of scheduling and for the non-pecuniary index.

Table 6: Effect of HHI on working conditions - OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	skills	autonomy	support	stability	development	physical safety	psy safety	scheduling	commuting	flexibility	intensity	NPI
Log HHI	0.0359	-0.156	-0.132	-0.0671	0.144	0.0193	-0.108	-0.0810	-0.000460	0.0480	-0.0367	-0.101*
	(0.0604)	(0.169)	(0.118)	(0.103)	(0.127)	(0.190)	(0.135)	(0.116)	(0.00690)	(0.194)	(0.162)	(0.0510)
Female × Log HHI	0.0187	-0.00440	0.122	-0.00859	-0.0769	0.00549	0.00835	-0.0410	-0.00631	-0.0897	-0.207*	-0.0215
	(0.0442)	(0.0851)	(0.119)	(0.0385)	(0.0904)	(0.115)	(0.0701)	(0.0988)	(0.00815)	(0.102)	(0.0934)	(0.0346)
N	3334	3342	3342	3342	3338	3340	2720	3342	3308	3340	3342	2686
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable each of the eleven working condition indicators and the non-pecuniary index, as described in Section 3.3.2. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). All columns include individual fixed effects, commuting zone-occupation fixed effects, and controls for the number of children. Survey weights are used.

Table 7: Effect of HHI on working conditions - IV HHI instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	skills	autonomy	support	stability	development	physical safety	psy safety	scheduling	commuting	flexibility	intensity	NPI
Log HHI	0.341*	-0.258	0.113	-0.283	0.452	0.365	0.223	-0.156	0.0126	0.693	-0.262	0.0582
	(0.168)	(0.455)	(0.330)	(0.275)	(0.297)	(0.402)	(0.288)	(0.286)	(0.0147)	(0.508)	(0.363)	(0.135)
Female × Log HHI	-0.0108	-0.0140	0.148	-0.00674	-0.153	0.0303	-0.0595	-0.0219	-0.00942	-0.0856	-0.213	-0.0376
	(0.0498)	(0.1000)	(0.130)	(0.0481)	(0.109)	(0.136)	(0.0921)	(0.121)	(0.00995)	(0.137)	(0.120)	(0.0390)
N	3334	3342	3342	3342	3338	3340	2720	3342	3308	3340	3342	2686
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using as a dependent variable each of the eleven working condition indicators and the non-pecuniary index, as described in Section 3.3.2. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. All columns include individual fixed effects, commuting zone-occupation fixed effects, and controls for the number of children. Survey weights are used.

Table 8: Effect of HHI on working conditions - IV Log 1/F

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	skills	autonomy	support	stability	development	physical safety	psy safety	scheduling	commuting	flexibility	intensity	NPI
Log HHI	0.255	0.304	-0.234	-1.068**	0.596	1.064*	0.312	0.359	0.00619	0.307	-0.130	0.235
	(0.154)	(0.600)	(0.407)	(0.326)	(0.365)	(0.518)	(0.353)	(0.325)	(0.0245)	(0.611)	(0.459)	(0.166)
Female × Log HHI	-0.0453	-0.131	0.198	0.141	-0.283*	-0.114	-0.147	-0.209*	-0.0112	-0.0821	-0.297	-0.104**
	(0.0563)	(0.154)	(0.146)	(0.0757)	(0.133)	(0.190)	(0.100)	(0.0974)	(0.00823)	(0.173)	(0.166)	(0.0391)
N	3334	3342	3342	3342	3338	3340	2720	3342	3308	3340	3342	2686
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ × occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using as a dependent variable each of the eleven working condition indicators and the non-pecuniary index, as described in Section 3.3.2. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. All columns include individual fixed effects, commuting zone-occupation fixed effects, and controls for the number of children. Survey weights are used.

5.3 Hirings

We now consider the effect of labour market concentration on the share of women among new hires. It has been shown in Marinescu et al. (2021) that concentration negatively affects the number of new hires. Indeed, if the firms with higher bargaining power cannot decrease wages, e.g. because of high minimum wages like it is the case in France, their increased bargaining power can translate into a decrease in the number of new employees. As shown in section 5.1, this higher monopsonistic power can affect women more than men. Firms that become more selective because they have less competitors in the labour market may prioritise men, perceiving women as more likely to interrupt their careers or reduce their labour supply due to family responsibilities. To test this hypothesis, we compute a measure of HHI for new hires, i.e. we keep only individuals who did not work in an establishment of their current firm the year before. We also compute the share of women newly employed in each firm and each year. Tables 9 to 11 present the effect of this newly computed HHI on the share of women hired, controlling for establishment-year fixed effects and adding occupation, then commuting-zone occupation fixed effects. When using linear regression (Table 9), we find a negative relationship between concentration and the share of women employed, which becomes positive and significant when including female-specific - occupation fixed effects but which does not hold when considering male-specific commuting zones. However, as mentioned earlier, potential endogeneity issues can bias the results, it is thus important to use an instrument for concentration. Tables 10 and 11 show that both IV approaches suggest a negative relationship between labour market concentration and the share of women among new hires. According to these estimations, a 10% in the index of concentration would decrease the share of women by 0.04-0.1 percentage point.

Table 9: Effect of HHI on the share of women hired - OLS

	(1)	(2)	(3)	(5)	(4)	(6)
	share	share	share	share	share	share
Log HHI	-3.983***	-4.782***	0.147**	-3.981***	-4.813***	0.0923
	(0.0296)	(0.0335)	(0.0519)	(0.0289)	(0.0329)	(0.0516)
N	1718149	1718149	1701697	1690105	1690105	1677958
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ	No	Yes	No	No	No	No
Male CZ	No	No	No	No	Yes	No
Female CZ × occupation FE	No	No	Yes	No	No	No
Male CZ × occupation FE	No	No	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). All columns include establishment-by-year fixed effects.

Table 10: Effect of HHI on the share of women hired - IV HHI instrument

	(1) share	(2) share	(3) share	(5) share	(4) share	(6) share
Log HHI	-7.857***	-7.722***	-0.636**	-7.598***	-7.391***	-0.403*
	(0.0479)	(0.0463)	(0.216)	(0.0470)	(0.0449)	(0.202)
N	1718088	1718088	1701648	1690042	1690042	1677908
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ	No	Yes	No	No	No	No
Male CZ	No	No	No	No	Yes	No
Female CZ $\times$ occupation FE	No	No	Yes	No	No	No
Male CZ $\times$ occupation FE	No	No	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. All columns include establishment-by-year fixed effects.

Table 11: Effect of HHI on the share of women hired - IV Log 1/F

	(1) share	(2) share	(3) share	(5) share	(4) share	(6) share
Log HHI	-5.745***	-5.560***	-1.297***	-5.721***	-5.507***	-1.158***
	(0.0494)	(0.0470)	(0.257)	(0.0486)	(0.0458)	(0.252)
N	1718088	1718088	1701648	1690042	1690042	1677908
Establishment-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Female CZ	No	Yes	No	No	No	No
Male CZ	No	No	No	No	Yes	No
Female CZ $\times$ occupation FE	No	No	Yes	No	No	No
Male CZ $\times$ occupation FE	No	No	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. All columns include establishment-by-year fixed effects.

6 Economic mechanisms

In this section we investigate a mechanism trough which labour market concentration can affect more women than men. Women have traditionally been less geographically mobile than men because they tend to bear most of the housework and childcare-related responsibilities. In our data, we cannot distinguish between women who are single and those who are in couple, but we have information on the number and the date of birth of their children. As shown earlier, having children decreases the geographical mobility for both men and women, with a more pronounced impact on women. We thus decide to recompute our commuting zones using the INSEE algorithm distinguishing by parental status in addition to the gender. We estimate an augmented version of equation (11) using the new definition of local labour market and including an interacted term capturing the parental status, to see if the effect of concentration increases the gender gap more among parents than non-parents. The results are presented in Tables E.1 to E.3 in Appendix. The linear estimation results in Table E.1 shows than having children increases wages for men but decreases them for women. It suggests that labour market concentration affects fathers more negatively than non-fathers, while there is no difference between mothers and non-mothers (if we sum the coefficient associated to Parent × Log HHI and the coefficient associated with the triple interaction, we obtain a result very close to zero). This would mean that the gender gap increases more with concentration among non-parents than among parents. However when using an instrumental variable approach, the coefficient associated with the triple interaction becomes negative, not significant at the usual levels when using the first instrument (Table E.2), but significant at the 1% level when using the second one (Table E.3).

Note that employer discrimination may not concern only women who already have children, but also women who are seen "at risk" of having children in the near future, in particular childless women of childbearing age may be discriminated against on the labour market because they may start to have children, take parental leave, and reduce their labour supply. Also, the exhaustive employer-employee data that we use in this paper do not allow us to distinguish between individuals with or without children (only the panel sub-sample has this information).
We thus replicate also the analysis on new hires by computing the share of women of childbearing age (below 41 years old) among new hires in each firm. Employers who have high bargaining power are likely to discriminate particularly against young women because they are more likely to have children, take childcare leaves, or to reduce their hours of work because of family responsibilities. We present the results in Tables E.4 to E.6 in Appendix. We can see that with our most demanding specification, a 10% increase in the index of concentration decreases the share of young women employed by 0.016 percentage points in the OLS estimation, but up to 0.3 - 0.7 percentage points with the IV estimations. These results suggest that women of childbearing age are particularly impacted by labour market concentration, and that parental status is a key mechanism in the differential impact of concentration across genders.

7 Conclusion

Monopsony power can have detrimental effects on workers, in particular through the worsening of pecuniary and non-pecuniary working conditions and the decrease in the demand for labour. Workers who are less geographically mobile, in particular women, may be even more affected because they can put less employers into competition. In this paper, we use a new definition of

commuting zones that are gender-specific and take into account these differences in mobility, to study how labour market concentration affects gender inequalities. We find that an increase in labour market concentration increases the gender gaps in hourly wages, and in some specifications, we also find that it worsens the scheduling possibilities, and decreases development opportunities and the working conditions on average. The negative impact on hourly wages is significant for both genders, but is between 30% and 100% higher for women than it is for men. We also find that concentration decreases the share of women among new hires, suggesting that employers who cannot act on wages instead decrease the demand for female workers.
We investigate the mechanisms behind these results by computing commuting zone that depend both on gender and on parental status, and by replicating the analysis to see if women with children, or women of childbearing age, drive these results. Indeed, their geographical mobility is lower, and they are often seen as less committed to work and more likely to interrupt their careers. Using instrument variables we find that the effect of concentration on the gender wage gap is higher among parents than non-parents, although this is significant only using of our instrument but not the other. However, we find that concentration significantly decreases the proportion of women of childbearing age among new hires, with a greater magnitude than when taking all women together.

Policy interventions limiting labour market concentration, such as anti-trust regulations, are likely to improve not only overall labour market outcomes but also mitigate gender inequalities. With regard to differences in geographical mobility, the growing availability of remote work may offer a solution for certain types of occupations. Lastly, if, as our results suggest, these results are driven by discrimination against women of childbearing age, the provision of accessible and affordable childcare options, or a more equal sharing of family responsibilities within households (such as trough more equal parental-leave taking) may eventually alter employers' perceptions of the risks associated with hiring women over men, even if not immediately. This could also have an impact on the differences in mobility among women (mothers) and men (fathers).
An interesting avenue for future research would be to focus on job-seekers and see how the duration of unemployment is affected by labour market concentration depending on characteristics such as gender, parental status and/or age group. One could also look at heterogeneous effects according to sectors/occupations, typically those who are more male or female dominated, or those that can offer more flexibility in terms of place of work or work hours.

References

Arquie, A. and Bertin, J. (2023). The heterogenous effects of employer concentration on wages: better sorting or uneven rent extraction?

Autor, D., Dorn, D., Katz, L., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of super start firms. *The Quartely Journal Of Economics*, 3:645–709.

Azar, J., Marinescu, I., and Steinbaum, M. (2022). Labor market concentration. *Journal of Human Ressources*, 57(s):S167–S199.

Barkai, S. (2020). Declining labor snd capital shares. *The Journal of Finance*, LXXV(5):2421–24–63.

Bassanini, A., Batut, C., and Caroli, E. (2021). Labor Market Concentration and Stayers' Wages: Evidence from France. *IZA Working Paper*, (14912).

Bassanini, A., Bovini, G., Caroli, E., Casanova Ferrando, J., Cingano, F., Falco, P., Felgueroso, F., Jansen, M., Martins, P., Melo, A., Oberfichtner, M., and Popp, M. (2023). Labour Market Concentration, Wages and Job Security in Europe. Working Paper Series No. 654, Nova SBE.

Becker, S. O., Fernandes, A., and Weichselbaumer, D. (2019). Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *Labour Economics*, 59:139–152. Special Issue on "European Association of Labour Economists, 30th annual conference, Lyon, France, 13-15 September 2018.

Benmelech, E., Bergman, N., and Kim, H. (2022). Strong employers and weak employees how does employer concentration affect wages? *Journal of Human Resources*, 57(s):S200–S250.

Bhaskar, V. and To, T. (1999). Minimum wages for ronald mcdonald monopsonies: A theory of monopsonistic competition. *Economic Journal*, 109:190–203.

Coombes, M. G. and Bond, S. (2008). *Travel-to-Work Areas: The 2007 Review.* Office of National Statistics, London.

Coombes, M. G., Green, A. E., and Openshaw, S. (1986). An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-to-Work Areas Revision in Britain. *The Journal of the Operational Research Society*, 37(10):943–953.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomc implications. *The Quarterly Journal of Economics*, 135(2):561–644.

Dube, A., Jacobs, J., S.Naidu, and S.Suri (2020). Monopsony in online labor markets. *American Economic Review: Insights*, 2(1):33–46.

Falch, T. (210). The elasticity of labor supply at the establishment level. *Journal of Labor Economics*, 28:237–266.

Franconi, L., D'Aló, M., and Ichim, D. (2016). Istat implementation of the algorithm to develop labour market areas. Technical report, Istat, Roma.

Franconi, L., Ichim, D., D'Alò, M., and Cruciani, S. (2017). Guidelines for labour market area delineation process: From definition to dissemination. Technical report, Istat, Italian National Institute of Statistics.

Gautier, P. and Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics*, 67:392–403.

Hirsch, B. and Janhn, E. (2015). Is there monopsonistic discrimination against immigrants. *ILR Review*, 68(3):501–528.

Hirsch, B., Schank, T., and Schnabel, C. (2010). Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from germany. *Journal of Labor Economics*, 28(2):291–330.

Jacob, N., Munford, L., Rice, N., and Roberts, J. (2019). The disutility of commuting? The effect of gender and local labor markets. *Regional Science and Urban Economics*, 77:264–275.

Jarosch, G., Nimczik, J., and Sorkin, I. (2021). Granular search, market structure, and wages.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender Differences in Job Search: Trading off Commute against Wage. *The Quarterly Journal of Economics*, 136(1):381–426.

Marinescu, I., Ouss, I., and Pape, L. (2021). Wages, hires, and labor market concentration. *Journal of Economic Behavior and Organization*, 184:506–605.

Matsudaira, J. (2014). Monopsony in the low-wage labor market? evidence from minimum nurse staffing regulations. *Review of Economics and Statistics*, 96(1):92–102.

Meiselbach, M., Eisenberg, M., Bai, G., Sen, A., and Anderson, G. (2022). Labor market concentration and worker contributions to health insurance premiums. *Medical Care Research and Review*, 79(2):198–206.

Moreno Galbis, E. (2020). Differences in working conditions between natives and immigrants: preferences vs. outside employment opportunities. *European Economic Review*, 130:103586.

Qiu, Y. and Sojourner, A. (2023). Labor-market concentration and labor compensation. *Industrial and Labor Relations Review*, 76(3):475–503.

Ransom, M. and Oaxaca, R. (2010). New market power models and sex differences in pay. *Journal of Labor Economics*, 28(2):267–289.

Rinz, K. (2022). Labor market concentration, earnings, and inequality. *Journal of Human Resources*, 57(s):S251–S283.

Robinson, J. (1969). *The economics of Imperfect Competition*. Macmillan.

Salop, S. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics*, 10:141–156.

Staiger, D., Spetz, J., and Phibbs, C. (2010). Is there monopsony in the labor market? evidence from a natural experiment. *Journal of Labor Economics*, 28(2):211–236.

Sullivan, D. (1989). Monopsony power in the market for nurses. *Journal of Law and Economics*, 32:S135–S178.

Webber, D. (2016). Firm-level monopsony and the gender pay gap. *Industrial relations*, 55:323–345.

Xiao, P. (2021). Wage and employment discrimination by gender in labor market equilibrium.

A Derivation of the steady state equilibrium conditions

Labour demand

Firms open vacancies until no more profit can be obtained so that equation (3) becomes then:

$$J_k = \frac{y - w_k}{r + \delta} \tag{13}$$

Combining this equation with equation (4) in the presence of the free entry condition leads to:

$$\frac{\gamma\theta(r+\delta)}{2m(\theta)} = \frac{u_F(y - w_F)\hat{x}_F + u_M(y - w_M)\hat{x}_M}{u_F + u_M} \tag{14}$$

Wage determination

Firms do not observe workers' most preferred geographical location or more precisely their x. As in Gautier and Zenou (2010), workers and firms only bargain over observable factors. Firms do not observe x but they observe $\hat{x}_k$ for $k = M, F$. As a result, in the wage bargaining, the value of employment is not given by (2) but by:

$$rE_k^d = w_k - \delta(E_k^d - U_k^d) \tag{15}$$

where U_k^d represents the disagreement payoff for the unemployed worker during the bargaining and is still given by (1). The match surplus of workers that is relevant for the wage bargaining is thus $(E_k^d - U_k^d)$ while that of firms equals $(J_k - V)$. The total surplus of the match is then given by: $\Omega_k = \left[E_k^d - U_k^d + J_k - V\right]$. Following the Nash bargaining process, the surplus is shared in constant proportions according to the respective bargaining powers of workers and firms. Denoting by $0 < \eta < 0$ the bargaining power of the worker, the sharing rule once we take into account the free entry condition $V = 0$, equals:

$$E_k^d - U_k^d = \eta\Omega_k \;\text{ and }\; J_k = (1-\eta)\Omega_k \;\;\Rightarrow\;\; (1-\eta)(E_k^d - U_k^d) = \eta J_k \tag{16}$$

From equation (1) we can easily find that $E_k^d - U_k^d = \frac{w_k - R}{\delta + 2m(\theta)\hat{x}_k}$. Replacing this expression and equation (13) in the Nash bargaining condition leads to:

$$w_k = \eta y \frac{(r + \delta + 2m(\theta)\hat{x}_k)}{r + \delta + \eta 2m(\theta)\hat{x}_k} + (1-\eta)R\frac{(r+\delta)}{r + \delta + \eta 2m(\theta)\hat{x}_k} \tag{17}$$

where $\frac{\partial w_k}{\partial R} > 0$ and $\frac{\partial w_k}{\partial \hat{x}_k} > 0$.[16] The higher the unearned income, the higher the wage. Similarly, the larger the range of acceptable geographical locations $\hat{x}_k$, the better the outside options and therefore the higher the bargained wages. Since θ, δ, y, R and r are the same for males and females, if $\hat{x}_M > \hat{x}_F$ males should be able to bargain higher wages.

[16] $\frac{\partial w_k}{\partial R} = (1-\eta)\frac{r+\delta}{r+\delta+\eta 2m(\theta)\hat{x}_k} > 0$ and $\frac{\partial w_k}{\partial \hat{x}_k} = \frac{2m(\theta)\eta(1-\eta)(r+\delta)(y-R)}{(r+\delta+\eta 2m(\theta)\hat{x}_k)^2} > 0$.

Range of acceptable geographical locations

We must finally determine the maximum distance a worker is willing to accept between her most preferred geographical location and the geographical location actually proposed by the job, *i.e.* $\hat{x}_k$. Formally, $\hat{x}_k$ is implicitly defined by the geographical distance that makes the worker indifferent between being employed or remaining unemployed:

$$E_k(\hat{x}_k, w_k) = U_k \tag{18}$$

Since employment for type-k depends linearly on x we can write $\frac{1}{\hat{x}_k}\int_0^{\hat{x}_k} E_k(x, w_k)dx = \mathbb{E}_x[E_k(x, w_k)|x < \hat{x}_k] = E_k(\mathbb{E}_x(x|x < \hat{x}_k), w_k) = E_k(\hat{x}_k/2)$ where the last steps follows from $\mathbb{E}_x(x|x < \hat{x}_k) = \hat{x}_k/2$. Therefore $\int_0^{\hat{x}_k} E_k(x, w_k)dx = \hat{x}_k E_k(\hat{x}_k/2)$. The unemployment asset becomes then:

$$U_k = \frac{1}{r + 2m(\theta)\hat{x}_k}\Big(R + 2m(\theta)\hat{x}_k E_k(\hat{x}_k/2)\Big) \tag{19}$$

Evaluating the employment asset value at $\hat{x}_k/2$ leads to:

$$E_k(\hat{x}_k/2, w_k) = \frac{w_k - \tau(\hat{x}_k/2) + \delta U_k}{r + \delta} \tag{20}$$

Replacing equation (20) in equation (19), and defining $\tau_k(x) = \mu_k x$, we can rewrite the asset value of unemployment as a function of $\hat{x}_k$:

$$U_k = \frac{1}{r + \delta + 2m(\theta)\hat{x}_k}\Big(R(r + \delta) + 2m(\theta)\hat{x}_k(w_k - \mu_k\hat{x}_k/2)\Big) \tag{21}$$

Evaluating the asset value of employment at $\hat{x}_k$, *i.e.* $E(\hat{x}_k, w_k) = \frac{1}{r+\delta}(w_k - \mu_k\hat{x}_k + \delta U_k)$, and replacing in (18) leads to:

$$(r + \delta)E(\hat{x}_k, w_k) \;=\; (r + \delta)U_k \Rightarrow w_k - \mu_k\hat{x}_k = rU_k \tag{22}$$

Replacing in the previous expression equation (21) we find:

$$0 \;=\; \hat{x}_k^2\mu_k m(\theta) + \hat{x}_k\mu_k(r + \delta) + (r + \delta)(R - w_k) \tag{23}$$

The strictly positive solution of equation (23) equals:

$$\hat{x}_k = -\frac{r + \delta}{2m(\theta)} + \frac{1}{2m(\theta)}\sqrt{(r + \delta)^2 + \frac{4m(\theta)(r + \delta)(w_k - R)}{\mu_k}} \tag{24}$$

B Definition of gender-specific local labour markets areas

B.1 The algorithm used to define labour market areas

The algorithm implemented in the R package *LabourMarketAreas* is an iterative agglomerative algorithm, based on Coombes and Bond (2008), that depends on a set of parameters. These parameters set the level of desired size and self-containment of the labour market areas (LMA thereafter). Self-containment can be imagined in terms of commuters who remain in the area. For example, a self-containment value of 0.75 means that the workforce who lives and works in the area is more than 75% of the commuters who only live in the area and more than 75% of those who only work in the area. The concept of self-containment of incoming and outgoing flows allows to quantify the most distinctive characteristic of a LMA, which is the ability to maximize the relationships inside its borders and minimize them across borders (Franconi et al., 2016). Let f_{hk} be the flow between municipality h and municipality k, *i.e.* the number of commuters living in h and working in k. Then, $R_i = \sum_k f_{ik}$ is the number of workers living in area i, $W_i = \sum_h f_{hi}$ is the number of workers working in area i and $RW_i = f_{ii}$ is the number of workers living and working in area i. There are two types of self-containment:

- the supply side self-containment: $SS\_SC = RW_i/R_i$
- the demand side self-containment: $DS\_SC = RW_i/W_i$

These two quantities measure the level of internal cohesion or integration of the areas with respect to the commuting flows (Franconi et al., 2016).

To be considered a LMA, a cluster of municipalities must have some minimum characteristics in terms of size and self-containment. The minimum size and minimum level of self-containment of a LMA may vary from country to another, but also from one region to another within a country, depending on the density of the population, the territorial morphology and the structure of commuting.

The algorithm allows the level of self-containment to change according to the size of the cluster so that it can be considered a LMA. This trade-off between size and self-containment is expressed by four parameters, defined by Coombes and Bond (2008), corresponding to target and minimum values of these characteristics:

- minSZ: minimum size of a cluster to be considered a LMA,
- tarSC: level of self-containment which is necessary for a cluster with minimum size to be considered a LMA,
- tarSZ: size of a cluster for which the minimum level of self-containment (minSC) is adequate for the cluster to be considered a LMA,
- minSC: minimum level of self-containment for a cluster that has size of at least tarSZ to be considered a LMA

The algorithm starts by considering each municipality as a cluster that is checked against a set of conditions to see whether it can be considered a LMA. At each stage, municipalities (or groups of municipalities aggregated previously) are aggregated according to the intensity of home-work exchanges. At each iteration clusters that are not fit for the purpose are disaggregated and a single municipality inside the cluster is chosen to be attached to a new cluster, improving the set of given conditions. The final solution is obtained when the whole set of clusters satisfies the given conditions (Franconi et al., 2016).

A validity condition, based on the parameter values, establishes the criteria that should be met by a cluster to be considered a LMA and quantifies whether the identified cluster is a valid LMA. This condition is operatively defined through a function that expresses the trade-off between the dimension (SZ), in terms of workers, and the self-containment (SC) of the cluster. This validity function, f_ν, depends also on the selected parameters and takes the following form:

$$f_\nu(SZ, SC) = \left[1 - \left(1 - \frac{minSC}{tarSC}\right) \cdot \max\left(\frac{tarSZ - SZ}{tarSZ - minSZ}, 0\right)\right] \cdot \left[\frac{\min(SC, tarSC)}{tarSC}\right] \quad (25)$$

The validity condition states that a cluster with size SZ_c and self-containment SC_c (minimum between SS_SC_c and DS_SC_c) is a proper LMA if:

$$f_\nu(SZ_c, SC_c) \geqslant \frac{minSC}{tarSC} \quad (26)$$

This condition is evaluated at each iteration to check whether all the clusters are indeed proper LMAs.

B.2 Implementation of the algorithm

In France, the values of these parameters have been defined at the national level, except for Corsica, Ile-de-France and the overseas departments, for which specific values have been defined. National parameter values in France are: minSZ=15 000, tarSZ=25 000, minSC=0.6, tarSC=0.7. The self-containment parameters are lower in Ile-de-France. In the case of Ile-de-France, a final choice was made to deviate slightly from the results of the algorithm, to respect the current limits of public establishments for inter-municipal cooperation (EPCI). Paris exerts a strong attraction on the surrounding areas, so that the other employment areas in Ile-de-France generally have a lower level of self-containment than that generally observed at the national level. Some LMAs in Ile-de-France have a self-containment level of less than 40% (*e.g.* Meaux, Etampes, Rambouillet). For the others, the self-containment level remains relatively low, between 41% and 57%, except for

Paris (89%). Even outside Paris, the size contrasts between labour market areas are significant, from 15 300 jobs for Provins to 336 000 for Roissy.

The analysis of the output of the package could reveal clusters which do not reach the minimum size required to define a cluster as a proper LMA or municipalities belonging to the reserve list (see Franconi et al., 2016), that were not assigned by the algorithms. All these cases can be classified under the heading of "self-contained cluster", *i.e.* a cluster which is completely self-contained (no flows inward and/or outward). This could be the case of a small island or groups of islands that does not reach the given threshold on the size or of remote municipalities. A manual assignment resolves this type of situations (*e.g.* islands are assigned to the labour market area where the connection with the mainland exists).

The zero list contains municipalities that could not be processed by the algorithm for various reasons: either the number of residents is 0 or the number of workers/jobs is 0 or the municipality has no interaction with any other municipality. In such cases, the algorithm eliminates the municipalities from the initial list (and let the user the choice to allocate them at a later stage). For mainland France outside Ile-de-France and Corsica, the zero list is composed of 1 041 municipalities, among which:

- 868 municipalities have a non-zero number of residents but zero workers
- 63 municipalities have a non-zero number of workers but zero residents
- 110 municipalities have (presumably) no interaction with any other municipality (non-zero number of residents and workers)

The algorithm of Coombes and Bond (2008) does not take into account in its production process any territorial contiguity principle. Therefore, areas that are non-contiguous might belong to the same labour market area. These need to be treated in order to create proper areas. This treatment, called fine-tuning, treats the non-contiguity by assigning part of the territory chosen by the user to one of the other nearby labour market area based on the function of cohesion used (see Franconi et al., 2016). There are different causes that may create non-contiguities, but some of them cannot be treated as they present structural characteristics of the territory and cannot be solved via algorithms/fine-tuning of the result (Franconi et al., 2017).

B.3 Parent-specific local labour market areas

We are able to define local labour market areas which are specific for parents (mothers and fathers) using the 2019 census data - professional mobility detail file. We identify parents using information on the structure of the household. Individuals who are parents in the data may belong to the following household structures: households whose main family is single-parent, households whose main family is a couple (two working persons, one working person and one inactive person, two inactive persons). Individuals who are not parents may be persons living alone or with other people without family ties (*e.g.* roommates). Unfortunately, we do not have precise information on the age of children living in the household. We only know the number of people in school in the household (including students) and the number of pupils, students or trainees aged 14 or over in the household. Therefore, we are not able to identify precisely the households with young children, where the labour supply is most constrained and women are likely to have a smaller job search area.

C Descriptive statistics

Table C.1: Average commuting distance between the municipality of residence and the municipality of work by gender and parental status

	(1)	(2)	(3)
	distance	distance	distance
Female	-11.24***	-11.06***	
	(0.0829)	(0.0823)	
Parent	-1.877***	-2.353***	-0.502***
	(0.0867)	(0.0865)	(0.140)
Female × Parent	0.397**	0.325*	-1.548***
	(0.131)	(0.130)	(0.214)
Intercept	33.456***	33.586***	28.115***
	(0.0553)	(0.0550)	(0.0482)
N	6308358	6302649	6187189
Year FE	No	Yes	Yes
Female CZ FE	No	Yes	Yes
Individual FE	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Panel Tout Salariés - EDP

The dependant variable is the distance between the municipality of residence and the municipality of work of the individuals in kilometers. It is computed by taking the distance between the centroids of the two municipalities.

Table C.2: Maximum commuting distance accepted by job-seekers by gender and parental status

	(1)	(2)	(3)
	max distance	max distance	max distance
Female	-4.566***	-4.538***	
	(0.0544)	(0.0539)	
Parent	-1.735***	-1.580***	0.0798
	(0.0543)	(0.0538)	(0.0880)
Female × Parent	-5.181***	-5.102***	-0.598***
	(0.0776)	(0.0768)	(0.129)
Intercept	29.670***	29.710***	26.934***
	(0.0335)	(0.0331)	(0.0337)
N	963056	963056	676296
Year FE	No	Yes	Yes
Female CZ FE	No	Yes	No
Individual FE	No	No	Yes

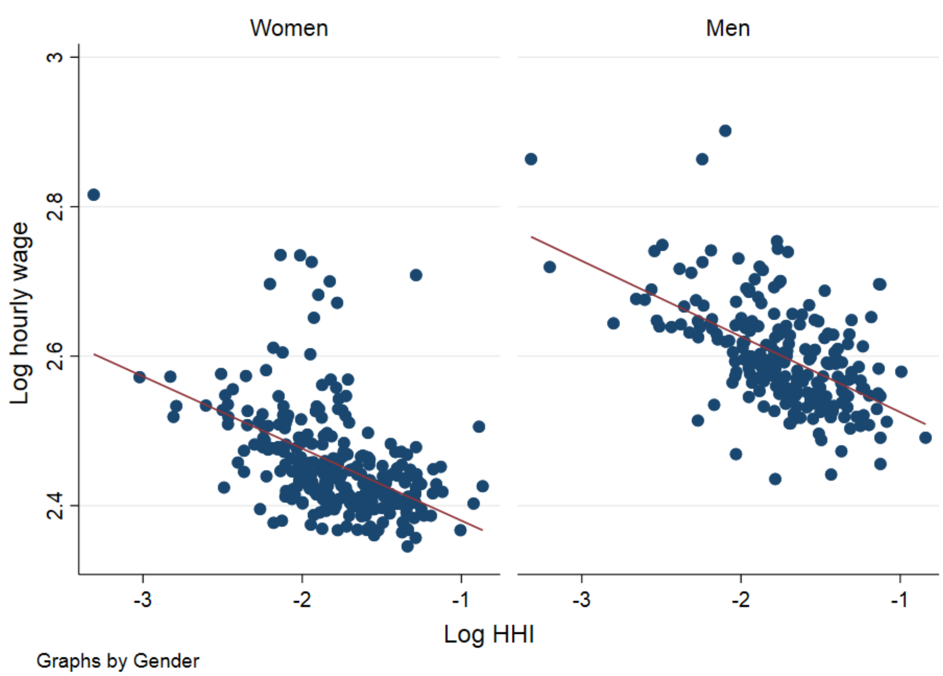
Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source : FH-DADS

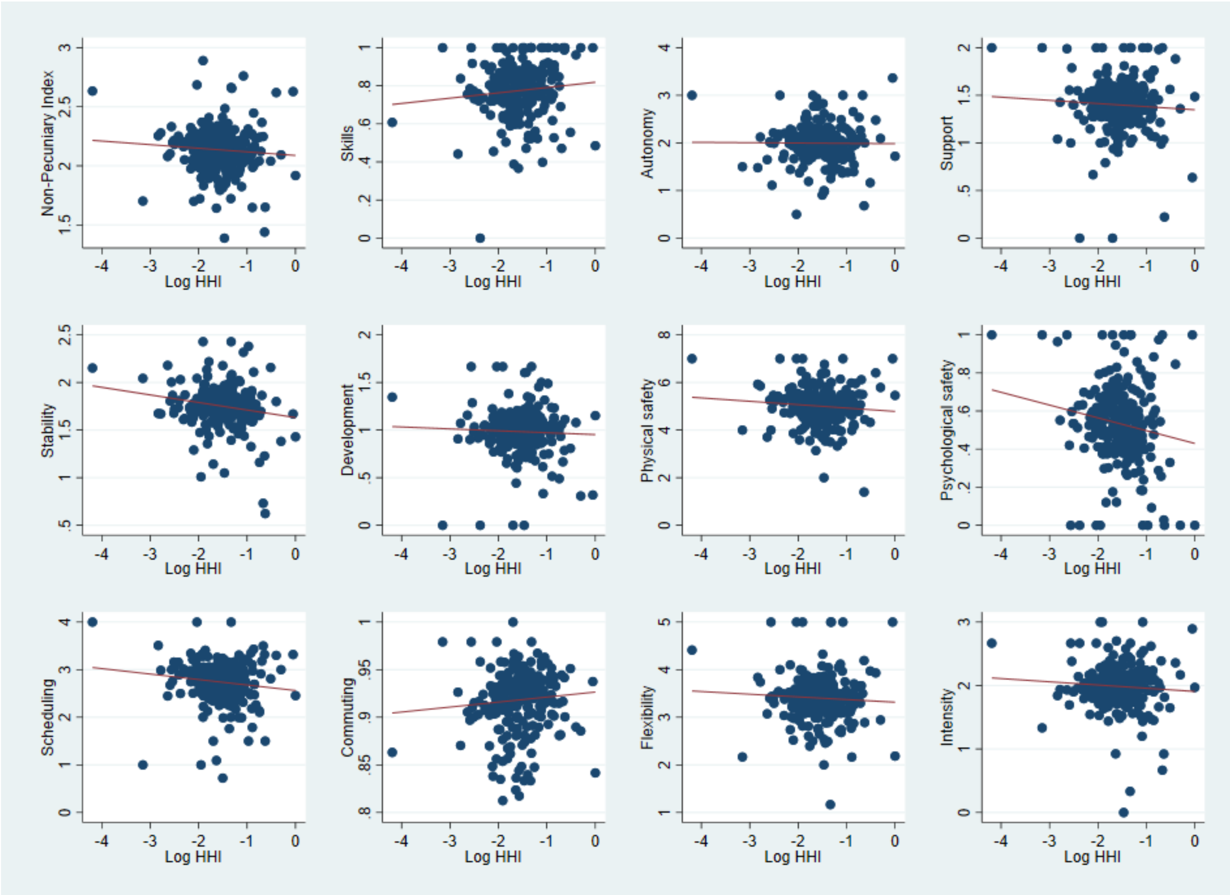
The dependant variable is the maximum distance that job-seekers are willing to commute in order to accept a job. It is given in kilometers. As there is only one observation per unemployment spell, the inclusion of individual fixed effects deletes individuals for which we observe only one unemployment spell. In this case, we exploit the difference in the willingness to commute of a given individual before and after he as a child.

Figure C.1: Relationship between log hourly wage and concentration by gender



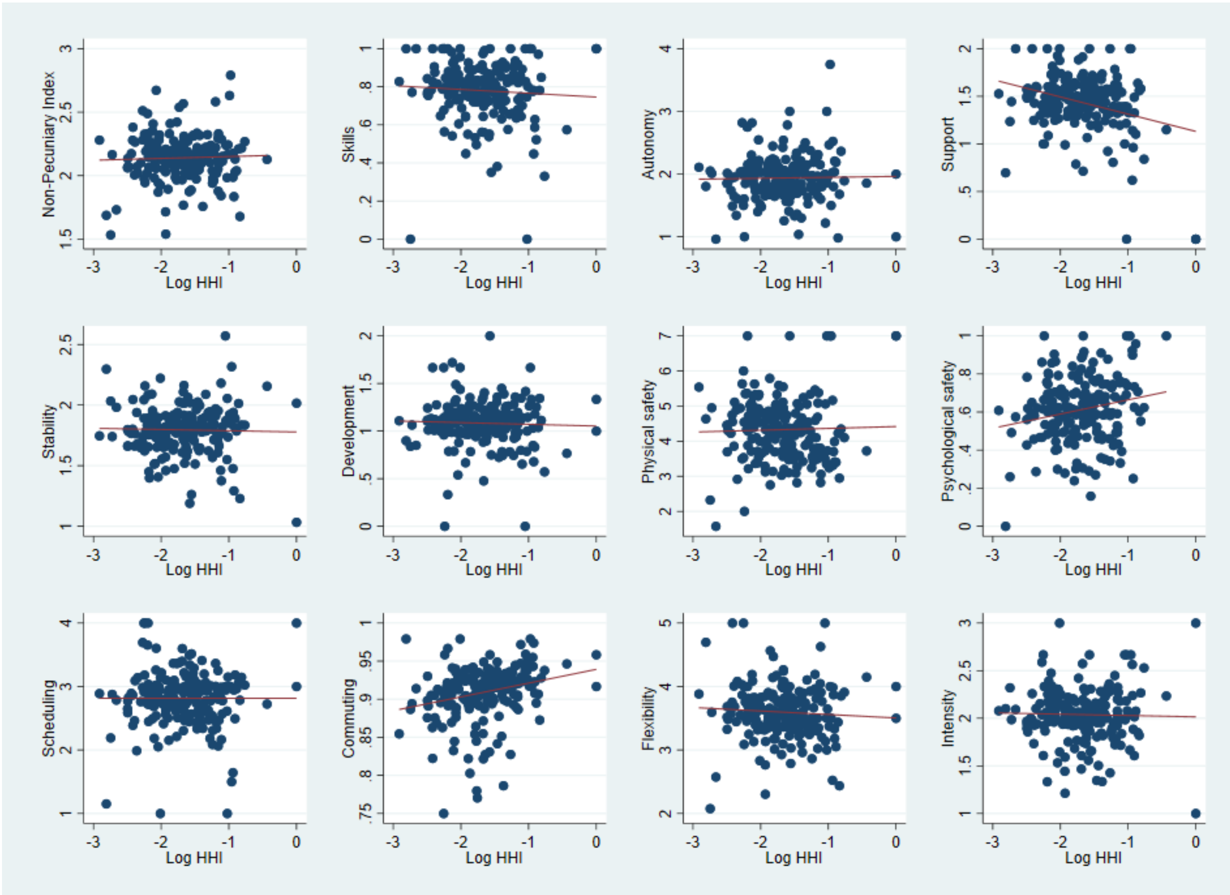
Source : Panel Tout Salariés - EDP

Figure C.2: Relationship between working conditions and concentration by gender - women



Source: Working Conditions Survey
Survey weights are used

Figure C.3: Relationship between working conditions and concentration by gender - men



Source: Working Conditions Survey
Survey weights are used

D Different measures of the HHI

D.1 Effect on wages - HHI based on new hires

Table D.1: Effect of HHI on hourly wages - OLS

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.102***	-0.0676***		
	(0.000980)	(0.000915)		
Log HHI	0.0405***	0.000304	0.000256	0.000166
	(0.000266)	(0.000399)	(0.000312)	(0.000313)
Female × Log HHI	0.000976**	-0.00149***	-0.00107**	-0.000979**
	(0.000314)	(0.000295)	(0.000365)	(0.000366)
N	2727016	2715801	2972271	2972271
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). It is computed on new hires only. Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience.

Table D.2: Effect of HHI on hourly wages - IV HHI instrument

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.125***	-0.0576***		
	(0.00165)	(0.00164)		
Log HHI	0.0297***	-0.00665***	-0.00127	-0.00221
	(0.000400)	(0.00174)	(0.00144)	(0.00145)
Female × Log HHI	0.00894***	-0.00494***	-0.00437***	-0.00326***
	(0.000558)	(0.000560)	(0.000790)	(0.000791)
N	2726973	2715763	2972222	2972222
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). It is computed on new hires only. To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience.

Table D.3: Effect of HHI on hourly wages - IV Log 1/F

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.0500***	-0.0160***		
	(0.00184)	(0.00192)		
Log HHI	0.101***	0.000638	0.00109	-0.000852
	(0.000422)	(0.00218)	(0.00195)	(0.00197)
Female × Log HHI	-0.0139***	-0.0195***	-0.00420***	-0.00409***
	(0.000627)	(0.000660)	(0.000810)	(0.000814)
N	2726973	2715763	2972222	2972222
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Female CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). It is computed on new hires only. To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. Controls variables include the number of children, the educational level, the work experience, the age, and a dummy variable for being born in France. When individual fixed effects are included, we keep only the number of children and work experience.

E Economic mechanisms: parental status

E.1 Hourly wages

Table E.1: Effect of HHI on hourly wages by parental status - OLS

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.111***	-0.0724***		
	(0.000600)	(0.000550)		
Parent	0.0418***	0.0316***	0.0306***	0.0323***
	(0.000672)	(0.000544)	(0.000749)	(0.000753)
Log HHI	0.0103***	-0.000910**	-0.000470	-0.000482
	(0.000277)	(0.000310)	(0.000329)	(0.000329)
Female × Log HHI	0.00627***	0.000825**	-0.000787*	-0.000815*
	(0.000294)	(0.000267)	(0.000359)	(0.000359)
Female × Parent	-0.0371***	-0.0192***	-0.0362***	-0.0399***
	(0.00101)	(0.000821)	(0.001161)	(0.001174)
Parent × Log HHI	-0.00532***	-0.00363***	-0.00105***	-0.00108***
	(0.000308)	(0.000252)	(0.000296)	(0.000296)
Female × Parent × Log HHI	0.00674***	0.00295***	0.00103*	0.00108*
	(0.000458)	(0.000372)	(0.000455)	(0.000455)
N	3666680	3627369	4022215	4022215
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Mother CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. Controls variables include the educational level, the work experience, the age, and a dummy variable for being born in France, in absence of individual fixed effects. The commuting zones used are the mother-specific commuting zones.

Table E.2: Effect of HHI on hourly wages by parental status - IV HHI instrument

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.110***	-0.0738***		
	(0.000676)	(0.000642)		
Parent	0.0449***	0.0312***	0.0319***	0.0336***
	(0.000829)	(0.000819)	(0.000934)	(0.000938)
Log HHI	0.0145***	-0.00221	-0.00266**	-0.00295**
	(0.000283)	(0.00126)	(0.000965)	(0.000967)
Female × Log HHI	0.00636***	0.00163***	-0.000817	-0.000747
	(0.000357)	(0.000334)	(0.000470)	(0.000470)
Female × Parent	-0.0511***	-0.0189***	-0.0317***	-0.0353***
	(0.00122)	(0.00116)	(0.00143)	(0.00144)
Parent × Log HHI	-0.00760***	-0.00317***	-0.00142***	-0.00145***
	(0.000407)	(0.000339)	(0.000400)	(0.000400)
Female × Parent × Log HHI	0.0132***	0.00277***	-0.000905	-0.000889
	(0.000591)	(0.000488)	(0.000600)	(0.000600)
N	3666678	3627368	4022214	4022214
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Mother CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from an instrumental variable regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). Controls variables include the educational level, the work experience, the age, and a dummy variable for being born in France, in absence of individual fixed effects. The commuting zones used are the mother-specific commuting zones.

Table E.3: Effect of HHI on hourly wages by parental status - IV Log 1/F

	(1)	(2)	(3)	(4)
	log(wage)	log(wage)	log(wage)	log(wage)
Female	-0.0794***	-0.0679***		
	(0.000739)	(0.000718)		
Parent	0.00362***	0.0176***	0.0308***	0.0328***
	(0.000942)	(0.00117)	(0.00106)	(0.00107)
Log HHI	0.0562***	0.0202***	-0.00361**	-0.00474***
	(0.000319)	(0.00216)	(0.00139)	(0.00140)
Female × Log HHI	-0.00752***	-0.00105**	-0.00137**	-0.00140**
	(0.000404)	(0.000383)	(0.000525)	(0.000525)
Female × Parent	-0.0238***	-0.00769***	-0.0272***	-0.0301***
	(0.00140)	(0.00152)	(0.00159)	(0.00160)
Parent × Log HHI	0.00797***	0.000745	-0.000715	-0.000765
	(0.000475)	(0.000418)	(0.000456)	(0.000456)
Female × Parent × Log HHI	-0.00399***	-0.00637***	-0.00295***	-0.00311***
	(0.000694)	(0.000579)	(0.000681)	(0.000681)
N	3666678	3627368	4022214	4022214
Controls	Yes	Yes	Yes	Yes
Establishment-year FE	Yes	Yes	Yes	Yes
Mother CZ × occupation FE	No	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Gender-year FE	No	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using the logarithm of hourly wages as a dependent variable. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. Controls variables include the educational level, the work experience, the age, and a dummy variable for being born in France, in absence of individual fixed effects. The commuting zones used are the mother-specific commuting zones.

E.2 Hirings

Table E.4: Effect of HHI on the share of women of childbearing age hired - OLS

	(1) share	(2) share	(3) share
Log HHI	-2.240***	-2.409***	-0.164**
	(0.0259)	(0.0290)	(0.0509)
N	1718836	1718836	1704586
Establishment-year FE	Yes	Yes	Yes
Female CZ	No	Yes	No
Female CZ × occupation FE	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women under 41 years old among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). All columns include establishment-by-year fixed effects.

Table E.5: Effect of HHI on the share of women of childbearing age hired - IV HHI instrument

	(1) share	(2) share	(3) share
Log HHI	-3.599***	-3.583***	-3.275***
	(0.0403)	(0.0394)	(0.210)
N	1718778	1718778	1704542
Establishment-year FE	Yes	Yes	Yes
Female CZ	No	Yes	No
Female CZ × occupation FE	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women under 41 years old among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the employment-weighted average of the HHI for the same occupation in the other commuting zones. All columns include establishment-by-year fixed effects.

Table E.6: Effect of HHI on the share of women of childbearing age hired - IV Log 1/F

	(1)	(2)	(3)
	share	share	share
Log HHI	-2.487***	-2.478***	-6.719***
	(0.0420)	(0.0403)	(0.249)
N	1718778	1718778	1704542
Establishment-year FE	Yes	Yes	Yes
Female CZ	No	Yes	No
Female CZ $\times$ occupation FE	No	No	Yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports results from a linear regression using as a dependent variable the share of women under 41 years old among the new hires, with a share ranging between 0 and 100. Log HHI corresponds to the logarithm of the concentration index computed at the intersection between a commuting zone and an occupation (with the HHI ranging between 0 and 100). To instrument Log HHI, we use the unweighted average of log(1/F) in other commuting zones, with F being the number of firms in the market. All columns include establishment-by-year fixed effects.